THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Essays in Financial Economics

FRANCESCO NICOLAI

A thesis submitted to the Department of Finance of the London School of Economics and Political Science for the degree of *Doctor of Philosophy*

June 2021

Declaration of Authorship

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Statement of conjoint work

I confirm that Chapter 2 is jointly co-authored with Simona Risteska. I contributed 50% of the work for Chapter 2.

I confirm that Chapter 3 is jointly co-authored with Marco Pelosi and Simona Risteska. I contributed 33% of the work for Chapter 3.

I declare that my thesis consists of 39,220 words.

Acknowledgements

I am extremely grateful to my advisor Dong Lou for his encouragement and inspiration throughout this endeavour. The numerous occasions on which I have reached out to him for advice have greatly helped me carry this dissertation through and grow as a researcher. I would like to extend my sincere gratitude to Ulf Axelson, Cameron Peng and Andrea Tamoni for their invaluable mentorship and support.

I also thank Ashwini Agrawal, Cynthia Balloch, Mike Burkart, Thummim Cho, Vicente Cuñat, Daniel Ferreira, Juanita Gonzalez-Uribe, Dirk Jenter, Christian Julliard, Paula Lopes-Cocco, Igor Makarov, Ian Martin, Martin Oehmke, Daniel Paravisini, Christopher Polk, Dimitri Vayanos, Michela Verardo, and many others for the countless insightful interactions that have made my time at the LSE so immensely instructive.

During my time at LSE I have certainly met some great friends who inspired me and made my PhD journey an experience to cherish. Among them special thanks go to Lukas Kremens and Marco Pelosi with whom I had the pleasure to share the *best row* in the office. I am grateful to Andreea Englezu, Zhongchen Hu, Bruce Iwadate, Karamfil Todorov, Yue Yuan, Agnese Carella, Fabrizio Core, Alberto Pellicioli, Lorenzo Bretscher, Juan Chen, Jesús Gorrin, James Guo, Brandon Han, Guido Maia, Olga Obizhaeva, Dimitris Papadimitriou, Bernardo Ricca, Gosia Ryduchowska, Petar Sabtchevsky, Amirabas Salarkia, Una Savic, Ran Shi, Arthur Taburet, Su Wang, Xiang Yin, for their insights and friendly chats.

I am very much indebted to the Paul Woolley Centre for the Study of Capital Market Disfunctionality at the LSE Financial Markets Group for financial support.

My deepest gratitude goes to my parents, Roberto and Donatella, for their love and support throughout my life, to Tamara for the great fun (and the contribution to Chapter 2!), and, mostly, to Simona for her enormous patience, love, tolerance and kindness. I have cherished every moment of this journey with you. This the-

sis is dedicated to my father who sadly passed away while I was working on this project; he was a kind, loving, bright man without whom I would not be the person I am. He will be greatly missed.

Abstract

This dissertation consists of three chapters. In the first, I study the impact of institutional investors on asset prices, by focusing on Collateralized Loan Obligations (CLOs). I document that, in order to satisfy constraints based on the par value of their assets, CLOs become forced sellers of leveraged loans. Loans sold for non fundamental reasons trade at depressed prices for up to nine months after the shock. The effect cannot be explained by selection on ex-ante or ex-post loan characteristics. A large fraction of the dislocation in secondary markets is transmitted to the market of issuance: shocked companies due to refinance their loans substitute away from institutional tranches towards other types of securities. I show that the substitution is imperfect, causing an increase in the cost of borrowing for affected firms.

In the second chapter, which is co-authored with Simona Risteska, we use data on mutual fund portfolio holdings to extract fund managers' stock return expectations. We use panel regressions and economic theory to demonstrate that we are able to partial out the effect of time-varying stock and manager characteristics (e.g., risk-aversion) and show that subjective expected returns are significantly affected by personal experience. Managers are more strongly influenced by recent returns and those experienced at the early stages of their holding period.

The third chapter, co-authored with Marco Pelosi and Simona Risteska, provides evidence of the disparity in the incidence of property taxes levied at different points in time. Housing demand is significantly less elastic with respect to taxes deferred to the future relative to taxes levied at the moment of the purchase. We attribute this difference to the lack of salience of future taxes at the moment of purchase. We provide directions on the optimal tax mix between salient and non-salient taxes with the help of a model.

Contents

List of Tables

List of Figures

1. Contagion in the Market for Leveraged Loans

FRANCESCO NICOLAI[1]

There is ample evidence that the supply of capital from institutional investors positively correlates with asset prices. It is however particularly challenging to determine whether institutional investors cause or respond to pricing conditions. This paper provides evidence in favour of the former by analysing the price impact of Collateralized Loan Obligations (CLOs) when they are forced to trade for non-fundamental reasons. CLOs provide an excellent setting to study the causal impact of institutional investors on the price of leveraged loans[2] both in their primary and secondary markets. The compensation of CLO managers is tested against the satisfaction of constraints based on the historical value of their assets. In order to avoid fire sales, these so-called *over-collateralization* (OC) constraints have been explicitly designed to be insensitive to the market price and the rating of leveraged loans as long as these securities are not impaired[3]. OC constraints have achieved their goal: CLOs have mostly avoided fire sales during the 2007-09 financial crisis and the market turmoil in March 2020 (Financial Stability Board, 2019; Kothari et al., 2020). However, historical-cost-based constraints might have unintended consequences: similarly to *gains trading* in the insurance industry (Ellul et al., 2015), when the

[1]I benefited from helpful comments by Ulf Axelson, Cynthia Balloch, Vicente Cunat, Daniel Ferreira, Dirk Jenter, Lukas Kremens, Dong Lou, Marco Pelosi, Cameron Peng, Simona Risteska, Karamfil Todorov, Yue Yuan and the seminar participants at LSE, the New Economic School, Universidad Carlos III de Madrid, BI Norwegian Business School, Tulane Freeman School of Business and Cass Business School. I am grateful to the Paul Woolley Centre for the Study of Capital Market Disfunctionality at the LSE Financial Markets Group for financial support.

[2]There is no standard definition of leveraged loans. Leveraged loans are usually structured and arranged by a syndicate of investment banks and they satisfy one or more of the following criteria: they are sub-investment grade securities rated at Ba1 (BB+) or below by Moody's (Standard & Poor's); they have a spread of at least 150bps over LIBOR; they are secured by a first or second lien against the issuer's assets (S&P Global, 2020b). The approach of this paper is to consider as leveraged loans the syndicated loans that are traded or held by CLOs.

[3]Leveraged loans are defined as impaired if their market price is below $85, their rating is at or below Caa (CCC) according to Moody's (Standard & Poor's), or they have defaulted (CreditFlux, 2015).

quality of some of their loans deteriorates, CLOs have an incentive to sell unrelated securities whenever they face binding constraints. A violation of these constraints would be extremely costly because it would force the management team to divert cash flows from subordinated to senior notes and would significantly decrease their compensation. The management team, therefore, engages in trading of securities that is fully aimed at avoiding CLO tests violations. This behaviour, in turn, generates contagion in the secondary market for leveraged loans: shocks spread from downgraded to otherwise healthy loans sold by distressed CLOs trying to fix their balance sheets. Consistently with the evidence in Loumioti and Vasvari (2019b) and the view of industry practitioners (Morningstar, 2018), I document that managers of CLOs that barely passed their OC tests tend to sell unrelated high quality loans compared to otherwise similar CLOs that failed them. In these regards I provide the following evidence: first, while OC constraints should mechanically deteriorate after a downgrade to Caa (CCC), I show that this is not the case, hinting at the fact that managers actively trade in order to restore their OC constraints after a downgrade. Second, CLOs are usually allowed to keep on their balance sheet up to 7.5% of Caa (CCC) or lower rated securities before their OC tests are impaired: I document that there is no bunching in the distribution of Caa (CCC) or lower rated securities at this threshold, suggesting that managers do not fix their OC constraints by trading downgraded loans. Third, I show that CLOs that have barely passed their OC tests sell securities that are trading at higher prices compared to CLOs that just missed their tests. This behaviour is particularly pronounced when the loans of a CLO have been downgraded to Caa (CCC). Fourth, loans sold by distressed CLOs that barely satisfy their OC constraints tend to have higher rating compared to those sold by CLOs that missed them.

I then proceed to analyse the impact on market prices of non-fundamental trades done by distressed CLOs to restore their constraints. I show that the price of loans sold by these CLOs is between 43.6bps and 71.8bps lower compared to otherwise similar loans. The effect is robust to the inclusion of various fixed effects which allows me to exclude selection based on ex-ante observable characteristics. The results are also robust to the inclusion of issuer-time fixed effects that are meant to compare the price of different loans issued by the same company, capturing the influence of any unobservable variable related to the fundamentals of the firm. I then test whether these trades are motivated by access to private information and refuse this hypothesis by showing that the loans sold by distressed CLOs to restore their OC tests are marginally less likely to default, and the likelihood of a downgrade following the trade is similar to the loans in the control group. This evidence allows me to conclude that the price change of loans sold by distressed CLOs is caused by their market impact rather than by their fundamentals. The price impact is long lasting: loans sold for non fundamental reasons trade at depressed prices for up to nine months after the event.

The effect of non fundamental trades is not restricted to the secondary market. I document that companies whose loans have been sold by distressed CLOs in the secondary market face a deterioration of their terms of financing when issuing new loans. I also provide evidence that the deterioration is caused by a reduction in the supply of funds from CLOs rather than by a reduction in companies' demand. I start by fixing the demand for capital along two dimensions: first I focus on firms that are due to refinance their debt in the following twelve months in the spirit of Almeida et al. (2011), and, second, I study the choice between institutional and bank tranches for firms that end up issuing leveraged loans by taking advantage of an identification strategy that is reminiscent of the one employed by Adrian et al. (2013), Becker and Ivashina (2014) and Fleckenstein et al. (2020). My findings are consistent with a reduction in the supply of institutional loans in the primary market: first, firms whose loans have been sold by distressed CLOs in the secondary market are between 3.9% and 11.7% less likely to include institutional tranches in their new issuances, compared to otherwise similar firms. Even when they are able to access the market for institutional tranches, these companies borrow between 5.5% and 11.3% less through this form of debt. Finally, the size of the tranches that are meant to be subscribed by CLOs is between 23.5% and 34.2% smaller compared to companies that did not suffer fire sales by constrained investors in the secondary market. The previous findings are paired with the evidence that treated companies face an increase in the cost of capital of up to 55bps, suggesting that indeed treated companies have to cope with a reduction in the supply of institutional capital from CLOs. How can we rationalize these findings? I show that CLOs are more likely to refinance in the primary market companies that they have previously invested in and held on their balance sheet. Companies whose loans are widely held and traded by distressed CLOs have a difficult time placing new loans that are issued at par and do not help alleviate the already tight constraints their usual investors are currently facing.

The rest of the paper proceeds as follows: Section 1.1 places the findings of this paper in the context of the existing literature; Section 1.2 explains the workings of CLO constraints, and, in particular, over-collateralization tests; Section 1.3 presents and describes the data used in the empirical analysis; Section 1.4 studies the trading behaviour of CLOs and how it is affected by OC constraints; Section 1.5 moves to the analysis of the price impact of CLOs in the secondary market for leveraged loans, documenting how loans sold by distressed CLOs trade at depressed prices; Section 1.6 proceeds to study how the distortions generated by CLO constraints propagate to primary markets by showing how the cost of capital and the financing decisions of companies are affected by shocks to distressed CLOs.

1.1 Existing Literature

The results briefly documented above add to various strands of the literature. First I contribute to the debate on the distortions generated by market-based and historical-cost-based constraints. In these regards, the closest papers to mine are Ellul et al. (2014) and Ellul et al. (2015) showing how historical cost accounting distorts the trading behaviour of insurers and spreads shocks across unrelated corporate bonds. CLO constraints, where the greatest majority of the assets are counted at their par value, are similar to historical cost accounting. As in Ellul et al. (2015), I show that a degree of insensitivity to fluctuations in market prices helps reducing the likelihood of fire sales, but leads to the spread of shocks from distressed to unrelated securities.

Second, I add to the literature on the propagation and amplification of shocks through financial intermediaries. Many have studied the impact of institutional investors on asset prices: in these regards, Shleifer and Vishny (2011) provide a comprehensive survey of the literature. Plenty of papers have been written on the relationship between mutual fund flows and asset returns both in equity and fixed income markets. Among others, Warther (1995), Wermers (1999), Nofsinger and Sias (1999), Coval and Stafford (2007), Frazzini and Lamont (2008), Lou (2012), Anton and Polk (2014) study the impact of mutual funds on equity markets. Schmidt et al. (2016), Chernenko and Sunderam (2016), Goldstein et al. (2017), Morris et al. (2017), and Zhu (forthcoming) focus on the impact on fixed income securities. A standard approach in the literature on mutual funds price pressure is to rely on shocks to flows as an instrument to study their price impact. However, it is particularly hard to disentangle whether flows cause or predict the subsequent movements in asset prices (Warther, 1995). On the other hand, the setting of the present paper presents a cleaner identification: the constraints on CLOs provide an incentive to trade unrelated securities compared to the ones that have been downgraded. These securities are often of good quality, guaranteeing that there is limited scope for reverse causality in my findings. Finally, it is relatively easy to control for ex-ante and ex-post measurable loan features allowing me to establish a chain of causation that goes from CLO forced trades to a deterioration in loan prices.

Third, I contribute to the literature on the trading and pricing of loans in the secondary market. Beyhaghi and Ehsani (2017) analyse the cross-sectional properties of loan returns in the secondary market; Gande and Saunders (2012), Allen and Gottesman (2006), Ivashina and Sun (2011b) and Addoum and Murfin (2020) study the relationship between loan prices and equity returns; Fabozzi et al. (2020) document inefficiencies in the secondary market for leveraged loans that are exploited by CLOs. I add to this literature by displaying how the prices for leveraged loans are affected by non fundamental CLO trades.

Fourth, this paper adds to the growing recent literature on CLOs. Benmelech et al. (2012) and Nadauld and Weisbach (2012) show that CLOs do not increase adverse selection in corporate loans and often lead to lower loan spreads; Bozanic et al. (2018) show that CLOs contribute to standardize the terms of loans contracts.

Loumioti and Vasvari (2019a) and Loumioti and Vasvari (2019b) are the first to study in detail the incentives generated by over-collateralization constraints on CLOs, documenting that tighter constraints lead to strategic trading aimed at overcoming tests based on them and, ultimately, cause a deterioration in performance. I confirm their findings, showing that CLO managers engage in loan trading that is ultimately motivated by the satisfaction of these constraints and use this feature to study the impact that CLOs have on the price of leveraged loans. Liebscher and Mählmann (2017), Fabozzi et al. (2020) and Cordell et al. (2020) provide evidence in favour of skill in CLO managers collateral selection and trading. Elkamhi and Nozawa (2021) analyse the trading behaviour of CLOs in periods of diffuse defaults or downgrades focusing on the effect that portfolio similarity across CLOs has on systematic risk. Fleckenstein et al. (2020) study the effect of CLOs on companies' financing outcomes. They use CLOs origination and repricing as an instrument for their appetite for new leveraged loans in the primary market and show that CLOs actively determine the amount of issuance in this market and are the main responsible for its cyclicality. This work adds on Fleckenstein et al. (2020)'s findings, with a particular focus on the heterogeneity of the impact of CLOs across companies: the firms whose loans have been sold by distressed CLOs fare worse outcomes independently of the aggregate appetite for leveraged loans by CLOs. My findings are also consistent with Allen and Gale (1994) who prove that, with endogenous market participation, relatively small idiosyncratic shocks can have large price effects on the underlying securities. The closest paper is the contemporaneous work by Kundu (2021) which replicates most of the findings in this article by employing a shift-share instrument and focusing on interest diversion rather than over-collateralization constraints.

Finally, I add to the abundant literature on the interaction between capital supply and companies' financing decisions. Among the others, Faulkender and Petersen (2006), Frank and Goyal (2009), Leary (2009), Sufi (2009), Lemmon and Roberts (2010). In the last part of the paper I employ an identification strategy that is similar to the one in Adrian et al. (2013) and Becker and Ivashina (2014), who show that firms with access to the bond market have shifted away from bank loans during the 2007-09 financial crisis, hinting at the fact that the credit collapse was due to a contraction in credit supply rather than demand. Ivashina and Sun (2011a) and Fleckenstein et al. (2020), on the other hand, study the substitution between institutional and bank tranches in the syndicated loan market; I expand on their findings by showing how issuer specific shocks in the secondary market force companies to rely on non institutional tranches at the expense of an increase in the cost of borrowing. This complements the literature on firms arbitraging capital market distortions[4]. I show that companies are able to only partially reduce the impact of shocks in the secondary markets by shifting from institutional to non-institutional loans, consistent with the idea of debt specialization (Colla et al., 2013).

1.2 The Mechanics of CLOs

A collateralized loan obligation (CLO) is a bankruptcy remote[5] investment vehicle whose purpose is to invest in fixed income assets, usually leveraged loans, and whose liabilities are represented by notes with decreasing seniority which is strictly enforced. Similarly to other types of securitizations, the interest and principal payments to noteholders come with a predetermined seniority, with senior notes receiving the first cashflows available to be distributed according to predetermined rules, followed by less senior notes and, eventually, equity holders. To enforce the seniority of the liabilities of a CLO, each tranche is subject to a battery of tests making sure that the priority of repayment is maintained. At each payment date CLO notes are tested against prespecified thresholds before the CLO is able to determine the distribution of cashflows to the noteholders. We can divide the tests into those that force the management team to divert cashflows from the junior to

[4]Loughran and Ritter (1995), Baker and Wurgler (2000), Dong et al. (2012) provide evidence on firms arbitraging equity markets misvaluations. Baker et al. (2003), Greenwood and Hanson (2013) and Harford et al. (2014) on firms timing debt markets. Gao and Lou (2013) and Ma (2019) on timing across different markets.

[5]Each party involved in the creation of a CLO, i.e., the originator, the arranger, the trustee and the manager, is separate from the assets which are placed in and legally held by a special purpose vehicle (SPV). The same applies to the CLO's liabilities. This guarantees protection from bankruptcy of the CLO for the parties involved and protection from bankruptcy of the parties involved for the owners of CLO's liabilities. This separation is highly sought after in order to guarantee that the creditworthiness of CLOs' liabilities is fully and uniquely determined by their assets. In order to obtain bankruptcy remoteness, the SPV must be a separate legal and operational entity, the assets must have been transferred via a *true sale*, and the originator cannot exercise control over the SPV (CreditFlux, 2015).

the senior notes, and the so called *maintain-or-improve* tests. The latter group - including tests on the weighted average rating factor (WARF) of assets, the weighted average spread (WAS), the industry concentration, and the weighted average life (WAL) of the collateral - are meant to reduce the riskiness of the CLO holdings, but have minimal impact on the trading behaviour of the management team (Loumioti and Vasvari, 2019b) because they tend to have ample slack and do not cause any diversion of cashflows from one class of notes to another. The other group, which is comprised of the over-collateralization (OC) and interest coverage (IC) tests, have more serious consequences for noteholders and have material consequences on the trading behaviour of the management team. Often OC tests are considered among the most onerous hurdles CLO managers have to face (CreditFlux, 2015). OC tests are a key tool in enforcing the seniority of the principal value of the notes of a CLO by guaranteeing that the value of the assets under management remains above a certain multiple of the par value of CLO tranches, while IC tests are meant to maintin the order of seniority of interest distributions. Given the paper mostly focuses on OC tests, I will describe their workings in greater detail. If we start from the most senior note (which I denote with the letter A) a CLO manager needs to make sure that the following constrain is not violated:

$$\widetilde{OC}^A \equiv \frac{\text{Par Value of Assets + Excess Cash}}{\text{Par value of Class A Tranche}} + \frac{\text{Market Value of Impaired Assets}}{\text{Par value of Class A Tranche}} \geq OC^A \tag{1.1}$$

where OC^A is a tranche specific threshold. It is obvious from equation (1.1) that, at the numerator of the test, assets under management are counted at par value (usually \$100 for leveraged loans in the United States) unless they have been impaired; in that case they might incur haircuts. The OC tests of CLO 2.0, i.e., the greatest majority of CLOs issued during and after the 2007-09 financial crisis, define as impaired securities all those assets that have defaulted or are at imminent risk of default, i.e., those rated at Caa1 (CCC resp.) or below by Moody's(Standard & Poor's)[6]. Defaulted securities are counted at the lower between their market value and the assumed recovery rate. The haircut for Caa (CCC) securities is more complex: each CLO has a Caa (CCC) bucket with a predetermined threshold, usu-

[6]Securities purchased at a deep discount, i.e. below $85-$80, are usually counted at their market value, as well.

ally 7.5% of the value of the CLO's collateral[7], and the loans above this limit with the lowest market value receive a haircut equal to the difference between their par value and their market price. Given that Caa (CCC) securities are generally traded at a significant discount compared to par, this guarantees that the numerator of the OC test is reduced whenever a CLO has more than 7.5% of securities rated at Caa (CCC) or below.

The same mechanism applies to any other tranche with lower seniority; for instance, for tranches with seniority equal to $k = B, C, ...$, the OC test can be represented in the following way:

$$\widetilde{OC}^k \equiv \frac{\text{Par Value of Assets + Excess Cash}}{\sum_{k' \leq k} \text{Par Value of Tranche } k'} + \frac{\text{Market Value of Impaired Assets}}{\sum_{k' \leq k} \text{Par Value of Tranche } k'} \geq OC^k \tag{1.2}$$

where the denominator sums over the par value of all the tranches k' with higher seniority than k. It is not an accident, but rather a featured design, that OC tests have little sensitivity with respect to swings in loan prices, given that, in their calculation, the greatest majority of assets are treated at their par value. The same holds true for any change in ratings that does not involve the Caa (CCC) bucket.

On the other hand, managers are particularly sensitive to changes in ratings that have a material effect on the Caa (CCC) bucket, given that breaching the limit of 7.5% forces managers to significantly reduce the value of their securities. Violating OC tests has serious consequences for a CLO manager: their compensation is, in fact, directly tied to the satisfaction of this metric. In particular, we can split the compensation of managers in three parts, two of which are conditional on the satisfaction of the OC test limit. The managing team first receives a *senior fee* which is a constant fraction of assets under management, usually 0.2%, and does not depend on the performance of the portfolio of loans. This fee is senior to any other payment to the noteholders and it is meant to be used to carry on with the daily management duties of the CLO. Second, we have the *junior fee*, which is earned after all the coupons to noteholders have been paid: this fee is approximately 0.3% of assets under management and it is meant to incentivize the management team

[7]Some deals, usually called "enhanced CLOs", are allowed to reach higher limits. They are however rare representing less than 0.3% of the market in 2019 (Goldfarb, 2019; S&P Global, 2019).

to make sure that the CLO's cashflows are sufficient to repay noteholders. Finally, we have the *incentive fee*. The incentive fee is paid after the equity holders of the CLO have achieved a pre-specified hurdle rate and consists of 20% of the cashflows above the prespecified hurdle.

Failing any of the OC tests (1.2) will reduce the compensation of the management team: any interest received must be diverted from paying notes with a lower seniority compared to the one that has failed the test; second, to make sure that the test is satisfied in the future, any principal repayment is diverted into repaying the principal of all the notes with a higher seniority; third, and most importantly, the management team does not receive neither the junior nor the incentive fee, de-facto cutting their income by more than half. It is therefore crucial for a manager to avoid breaching any OC test. While this high-powered incentive is designed to make sure that the interests of the management team are aligned with the interest of the liability holders, it might also be responsible for a distortion of the objective of the management team. It is an empirical question whether and how the threat of failing an OC test affect the trading behaviour of CLO managers: this specific issue is tackled in Section 1.4. Before delving into this topic, the next section introduces the data I use in the rest of the paper.

1.3 Data and Summary Statistics

In order to analyse the trading behaviour of CLO managers and their effect on leveraged loans, we need to collect data on the holdings and trades of CLOs, together with the characteristics of the underlying loans. All the data on CLOs come from the CreditFlux Clo-i dataset. CLOs are required to provide their investors with a quarterly payment report and a monthly trustee report which have been collected by CreditFlux starting from 2009. More specifically, I make use of granular data on the universe of CLOs' holdings at their reporting dates, all the transactions that CLOs have completed between reporting dates and all the results of the OC tests which constrain the CLOs in my sample. The focus of the paper is on US based CLOs in the period between January 2009 and December 2019. Overall I have access to 89,111 unique reports from 2,601 distinct CLO deals, supervised by 218 distinct managers. Summary statistics are provided in Table A1, where Panel A reports the statistics for each individual CLO deal at each reporting date, while Panel B presents data aggregated at the management team level. The average (median) CLO deal has $434.73Mln ($408.86Mln) assets under management, comprised by 213.82 (203) unique securities from 168.2 (157) distinct issuers. Each security, on average, represents 1% of the total portfolio. The average age for a CLO is 2.57 years, reaching up to 11.90 years. If we look at the characteristics of the securities held by CLOs, the median fraction of assets rated at Caa (CCC) or below is about 6%, with a mean of 9%. On the other hand the median fraction of defaulted assets is equal to 2%, while the median is equal to 5%.[8]. The typical team contemporaneously manages 5.57 CLOs for a total of $2.505Bln of assets. The average number of securities held is 565.32 issued by, on average, 257.45 distinct companies. Management teams are on average 4.70 years in sample.

I then proceed to analyse the transactions carried out by CLOs whose summary statistics are reported in Table A2. First of all, it should be noted that CLOs do not report a trade in roughly 18% of the months (i.e., they have zero trades for 16,596 reports out of 89,111 total reports). When they do trade, the average number of transactions is equal to 29.73 while the median number of transactions is 19. If we analyse purchases and sales separately, CLOs do not buy any security in 26.5% of the reports, while they do not sell any security in 26.8% of the reports. CLOs tend to purchase securities in the ramping-up period when they are trying to reach the desired amount of assets under management, while sales happen during the whole life of the deal: this implies that the average number of purchases is larger than the average number of sales (19.25 vs 13.69) given that securities are usually bought in bulk at the beginning of the life of a CLO. This is confirmed by the figures on the total amount of transactions: on average CLOs report purchases for $22.44Mln and sales for $10.84Mln. These figures are highly skewed with the largest amount purchased equal to $2.37Bln and the largest amount of securities sold equal to roughly

[8]Some CLOs end up having 100% of assets rated at CCC or below before they are shut down, skewing the mean towards larger values. The same holds true for defaulted securities.

$1Bln. The typical security is sold at a price of $98.20, while the typical purchase price is $98.99. There are two possible explanations for this divergence: either CLOs tend to sell worse performing and purchase better performing securities, or CLOs tend to purchase securities in periods of relative calm and become forced sellers in periods of distress. I will show in Section 1.4 that the latter explanation is the more plausible among the two: in periods of distress, when downgrades become more prevalent, CLO managers become forced sellers. Further statistics on CLO holdings and transactions are provided in Appendices A.4 and A.5.

Finally, the analysis in Section 1.6 focuses on the impact that forced sales have on the primary market for U.S. syndicated loans. The data on loan originations is collected from Refinitiv's Security Data Company (SDC) Platinum[9]. The dataset includes historical information on more than 315,000 global corporate loan transactions since the early 1980s. In the paper I will focus on loans issued by companies domiciled in the United States between January 2009 and December 2019. In order to compute the exposure of issuers to loan rollover I also include all the loans maturing after January 2009, but potentially issued before my sample period starts. The overall sample includes 76,610 unique tranches from 48,757 facilities issued by 19,378 distinct borrowers. No common identifier between the SDC Platinum and the CLO-i dataset exists; for this reason, Appendix A.3 outlines the matching procedure used to link the two datasets. The analysis in Section 1.6 makes use of the number of loans, the tranche size and the All-in Spread Drawn (AISD) obtained from SDC Platinum.

1.4 Trading Behaviour

In Section 1.2 I have outlined how the compensation of CLO managers is tied to the satisfaction of OC tests. In this section I analyse how managers' portfolio choices are affected by the incentives generated by OC tests and whether their trading behaviour responds to the deterioration of this metric. Equation (1.2) clearly shows that OC tests are affected by external shocks only when securities default or when

[9] Thomson Financial, I. (2001).

they are downgraded to Caa (CCC): in what follows, I focus on the second type of shock[10]. In particular, in the remaining of this section, I show the following: managers do actively trade in order to restore their OC constraints when hit by downgrades to Caa (CCC); because there is a certain degree of freedom in booking Caa (CCC) rated securities at an inflated value on the balance sheet of CLOs, managers avoid directly selling them; CLOs restore their OC tests by building par, namely by selling high quality securities that trade at or above par and buying low quality securities that trade well below par.

I start by providing evidence that managers actively trade in order to restore their OC constraints. When a security is downgraded to Caa (CCC), if the Caa (CCC) bucket is already above the 7.5% threshold, the marginal Caa (CCC) security receives a haircut which, in turn, impairs the OC test. This implies that, absent any active measure by the manager, OC tests should mechanically deteriorate when a CLO is hit by a downgrade. I therefore measure the slack of an OC test as the percentage distance from the predefined test's threshold, i.e. $\text{slack}_{i,t}^{k} = \frac{\widetilde{OC}_{i,t}^{k} - OC_{i}^{k}}{OC_{i}^{k}}$, where $\widetilde{OC}_{i,t}^{k}$ is the realization of the OC test for tranche k and CLO i, while OC_{i}^{k} is the predetermined test threshold. I then construct two dummy variables: $\text{Shocked}_{i,t}$ which takes a value of one whenever a CLO has been hit by downgrades to Caa (CCC), and Above $7.5\%_{i,t}$ which takes a value of one when the Caa (CCC) bucket is above the threshold of 7.5%. I proceed by regressing the percentage slack of a CLO's OC tests on these variables and report the results in Table A3, from which we can conclude that there is no mechanical relationship between tests and shocks. In particular, columns (1)-(4) show that, after having been hit by downgrades to Caa (CCC), the slack of OC tests is neither higher nor lower: these results are robust to the inclusion of time, CLO deals, and type of test (Junior vs Senior) fixed effects. When the interaction $\text{Shocked}_{i,t} \times \text{Above } 7.5\%_{i,t}$ is included in column (5), there is still no evidence that the slack of CLOs hit by downgrades to Caa (CCC) and whose Caa (CCC) buckets are above 7.5% is lower than average.

[10]Table A1 shows that defaulted securities tend to be a relatively small fraction of CLOs' collateral; moreover, defaulted securities are usually first downgraded to Caa (CCC): conditionally on being rated Caa (CCC), the probability of default within one year is about 25% (Elton et al., 2001; Fei et al., 2012; S&P Global, 2020a).

Recall that there should be a mechanical negative relationship between the slack of an OC test and the number of Caa (CCC) rated securities as shown in equation (1.2). The lack of any correlation between the two hints at the fact that managers respond to incentives and actively trade to restore their OC tests. When assets are downgraded to Caa (CCC) between reporting periods, managers actively trade before the next report to make sure their compensation is not reduced. As a result, the slack of CLOs that suffered a downgrade is no different from the slack of the ones that did not.

After having determined that managers respond to incentives and actively trade to restore OC tests, we are left wondering what actions can a manager undertake in order to improve this metric. There are only two ways to gain slack on the OC test: the first way is to directly get rid of Caa (CCC) securities or any security that is counted at market value, while the other is to *build par*. Selling Caa (CCC) loans that are recorded at their market value can help as long as the marginal security, i.e., the security that was crossing the 7.5% threshold, does nor receive a haircut and is recorded at par-value or if the proceeds are used to purchase securities that trade at a discount. This is, however, a relatively ineffective strategy for two reasons. First there is a purely quantitative consideration: Caa (CCC) loans trade at large discounts hence it takes a relatively large number of securities to gather a sizeable amount of cash. Second, selling a Caa (CCC) security would force the manager to recognize a capital loss that could possibly be avoided. This is because distressed loans usually receive a haircut that is in line with their market price if they have been recently traded; managers, however, can value distressed loans at a theoretical bid price obtained from a dealer in case there is no pricing information(CreditFlux, 2015). This price is then certified by the CLO trustee. Managers may have some freedom in recording Caa (CCC) loans at a higher price compared to the one they could realise by selling the security. In order to test this hypothesis I make use of the fact that, since September 2017, CLO-i reports data on the price at which each loan has been recorded on CLOs' monthly reports. I can therefore directly compare the price at which loans are recorded on the balance sheet of a

CLO with their closest market price by running the following regression:

$$\text{discount}_{j,t} = \alpha_j + \alpha_t + \beta_1 \text{Transaction}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.3}$$

$\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$ is the discount compared to par[11], $\text{Transaction}_{j,t}$ is an indicator variable equal to one whenever the price comes from an actual transaction (purchase or sale) by a CLO and zero otherwise, α_j and α_t are issuer and time fixed effects, while $X_{j,t}$ includes a set of fixed effects for rating, industry and interest rate of the loan. Table A4 reports the results of regression (1.3), with Panel A focusing on all loans, while Panel B restricting the attention to loans with a rating of Caa (CCC) and below. As expected, the results indicate that managers have a certain freedom in keeping loans at inflated prices on their balance sheet. After having controlled for the effect of confounders, the average transaction happens at a discount of 87bps compared to the average price at which the loan is recorded on the trustee report by CLO managers. The result is robust to the inclusion of Year×Month×Issuer fixed effects that are meant to partial out any unobservable issuer characteristics. The impact is much larger when we focus on Caa (CCC) loans in Panel B: the average loan rated at Caa (CCC) or below is usually traded at a discount of 400bps compared to the average price recorded on CLO monthly reports. Selling these securities is particularly costly in order to gain slack in the OC constraints.

CLOs, however, are not required to sell their entire bucket of Caa (CCC) securities: the management team needs to make sure that the number of Caa (CCC) loans does not exceed the threshold of 7.5%. This implies that it might be optimal for the CLO to guarantee that these securities never cross the 7.5% and we can use this fact to check whether managers actively trade these loans to pass the OC tests.This is the aim of Figure A1 plotting the empirical distribution of Caa (CCC) securities as a fraction of assets. The figure shows that there is no obvious discontinuity at 7.5%. I formally test the presence of such a discontinuity using Cattaneo et al. (2020) local polynomial density estimator. Table A5 provides no evidence in favour of such

[11]For a loan with price $P_{jt} = \$100 - \$X$ and par value of \$100, the discount is equal to $100 \times \log(\$100/(\$100 - \$X)) \approx X$ percent.

a discontinuity, suggesting managers do not actively keep the pool of Caa (CCC) securities at the 7.5% threshold.

If managers do not trade loans recorded at their market value, they can engage in *par building* in order to improve their OC tests. Par building involves a trade whereby a manager sells a highly quality security, possibly trading above par, in favour of a lower quality security, trading below par. For instance, a CLO in need of \$100 of par value can adopt the following strategy: sell nine loans that are currently trading at \$100; the transaction will generate proceeds for \$900 which the manager will immediately use to buy ten loans that are currently trading at \$90. In market value terms the transaction is neutral, given that the proceeds from the sale can be used to buy the new loans. However the OC test will improve: nine loans trading at par contributed for 9 × \$100 = \$900 in the OC test, while ten loans trading at \$90 will contribute for their full par value in the OC test,i.e., 10 × \$100 = \$1000 in the numerator of the OC test. The rest of the section will show that, indeed, managers resort to par building in order to improve their OC tests.

First, par building should be more aggressive for those CLOs whose OC constraints are close to be binding. In order to test this hypothesis, I divide CLOs in seven buckets[12] based on the slack of their OC tests. I measure the amount of par gained with each transaction as: $\text{gain}_{i,j,t} = 100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for purchases and $\text{gain}_{i,j,t} = -100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for sales, where $P_{j,t-1}$ is the last transaction price available for security j before being traded by CLO i. Notice that the lagged price $P_{j,t-1}$ is not affected by the CLO transaction and captures the price a manager faces when deciding which loans to sell in order to restore their OC tests. I then sum across all the transactions by a CLO in any given reporting period to construct the following variable: $\text{gain}_{i,t} = 100 \times \left(\sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$[13].

[12] The buckets are the following: [-100%,-5%), [-5%,0%),[0%, 5%), [5%,10%), [10%,15%),[15%,20%),[20%, 100%).

[13] To understand the construction of the gain, it might be instructive to look at a specific example. If a CLO with assets of \$10,000 sells nine loans trading at \$100 then the gain is equal to $-(\$100 - \$100) \times \frac{9}{\$10{,}000} = 0$, while if the manager buys ten loans trading at \$90 the gain is equal to $(\$100 - \$90) \times \frac{10}{\$10{,}000} = +0.009$. The two transactions have generated an increase in par value equal to $\frac{(\$100-\$100)\times 10+(100-90)\times 9}{\$10{,}000} = +0.009$. If we multiply by 100 we get the figure in percentage terms, i.e.

The left panel of Figure A2 plots the average amount of par gained as a function of the slack of Junior OC tests, while the right panel as a function of the slack of Senior tests. It is clear that CLOs whose Junior OC tests are binding, i.e. those with slack close to 0%, are more likely to engage in par building compared to any other CLOs. The large discontinuity between the CLOs that have just missed an OC test and those that have barely passed it suggests that par building is a key tool used by managers in order to guarantee the satisfaction of their constraint and is consistent with previous evidence by Loumioti and Vasvari (2019a,b). Figure A16 confirms that indeed there is a clear discontinuity in par gained between these two groups of CLOs. No such a discontinuity is present in Senior OC tests, signalling that these are less important in determining the trading behaviour of CLO managers.

After having performed this unconditional analysis we can study the behaviour of CLOs that have a binding constraint in the month when they are hit by downgrades to Caa (CCC). For this purpose I construct a dummy variable $\text{Shocked}_{i,t}$ that is equal to one whenever the loans of a CLO have been downgraded to Caa (CCC) and a dummy variable $\text{Constrained}_{i,t}$ that is equal to one whenever the slack of a CLO is between 0% and 5%, i.e., the CLO is in the group that we have shown to be more likely to engage in par building. Columns (1) and (2) in Table A6 report the results of the following regression:

$$\text{gain}_{i,j,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t} \tag{1.4}$$

while, columns (3) and (4) report the results of the following regressions, where I sum across all the transactions of a CLO at time t:

$$\text{gain}_{i,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t} \tag{1.5}$$

Columns (1) and (3) measure the slack in terms of Junior OC tests, while columns (2) and (4) in terms of Senior tests. From all the regressions we can infer that shocked CLOs are more likely to engage in par building: each transaction of a

0.9 percent.

shocked CLO contributes towards increasing par by between 0.005% and 0.006%, while when we sum across all the transaction carried out by a CLO, we can conclude that shocked CLOs build between 0.033% and 0.050% more par compared to the CLOs in the control group[14]. When we compare the differential effect of Junior and Senior tests, i.e. we contrast the results in column (1) and (3) with those in columns (2) and (4), we can conclude that only Junior tests matter for par building: CLOs whose Junior slack is between 0% and 5% build 0.035 more par per period compared to other CLOs, while being shocked adds an extra 0.069 in par. The overall effect for CLOs that are constrained and have suffered from downgrades is of 15.4bps (i.e., 3.5bps + 5bps + 6.9bps).

Similarly, we can analyse the probability of selling securities that trade above par or that are rated Caa (CCC) in the month other loans on the same balance sheet have been downgraded. Figure A3 and A4 show that CLOs that barely passed their OC tests are more likely to sell securities above par and less likely to sell securities that are rated Caa (CCC) or below in the month of a downgrade. It should be stressed that CLOs that barely passed their tests are likely to be different from those that missed them along many dimensions. However the scope of the paper is to study the impact of CLOs on the price of loans. Therefore even if the two groups of CLOs differ systematically, this does not invalidate the identification strategy employed to study the price impact of CLOs. We only need to prove that the loans sold by constrained CLOs are sold for non fundamental reasons. Even if the subset of securities held by the two groups were systematically different, we could control for this difference with the help of issuer fixed effects that would take care of this type of selection.

It should be pointed out that par building helps a manager in locking-in trading gains by selling well performing loans and buying loans that are trading at a lower price. However this does not represent a free-lunch: the loans that are sold might be significantly safer compared to the ones that are purchased, resulting in risk-shifting. It is intuitive that constrained and unconstrained CLOs might engage

[14]While these figures might seem small, notice that each transaction represents a small fraction the assets of a CLO. Moreover notice that the improvement in OC tests is equal to $\text{gain}_{i,t} \times \frac{\text{Principal Balance}_{i,t}}{\text{Par Value of Note } k_{i,t}}$ where $k =$ Junior, Senior.

in par building for different reasons: in normal times CLOs want to improve the quality of their collateral, while - when the constraint is binding - to gain slack on their OC tests. One way to test this hypothesis is to look at the riskiness of the loans that are sold and purchased by constrained and unconstrained CLOs: if indeed the trading behaviour of constrained CLOs is motivated by a desire to restore the strength of OC tests, then we should expect them to sell safer loans to finance the purchase of riskier ones. In order to investigate this hypothesis, I measure the riskiness of a loan with its rating factor. CLOs are, in fact, subject to tests on the Weighted-Average Rating Factor (WARF) of their assets, where loans are assigned a numeric value on a scale from 1 to 10,000 which is supposed to capture their riskiness. I make use of Moody's rating mapping (reported in Table A7) in order to transform the ordinal ratings into a meaningful cardinal value. The rating factor increases with the riskiness of a security, ranging from 1 (Aaa) to 10,000 (Ca-C). The typical rating for a loan on the balance sheet of a CLO is B2, commanding a factor of 2,720, while the average and median WARF are 2,923.48 and 2,737, respectively. As one should expect, WARFs tend to vary and follow the business cycle, as Figure A14 shows.

Table A8 formally tests whether constrained CLOs tend to sell higher quality securities by regressing the rating factor of loans sold and purchased on the $\text{Shocked}_{i,t}$ and $\text{Constrained}_{i,t}$ dummies. The table shows that unconstrained CLOs sell lower quality loans (average rating factor of 3025) compared to the ones they buy (average rating factor of 2641), implying that - in normal times - their trading decisions make their collateral safer. When CLOs are hit by downgrades, however, they sell loans of slightly higher quality and purchase loans of lower quality: on average, the assets they sell have a rating factor that is 151.6 points lower, while their purchases have a rating factor that is 67.5 points higher. The coefficient on the $\text{Constrained}_{i,t}$ dummy, suggests that, when they have not been hit by downgrades, the quality of the loans they sell is slightly worse (73.9 points), while there is no difference in the loans they purchase. Finally the coefficient of interest, i.e. on $\text{Shocked}_{i,t} \times \text{Constrained}_{i,t}$, shows that after having suffered a downgrade to Caa (CCC), CLOs whose constraints are binding sell loans that are significantly safer

(-514.2 points), while they purchase riskier loans (105.8 points). The results in the table suggest that the average rating factor of loans sold by distressed CLOs is 591.9 points (i.e., -151.6 + 73.9 - 514.2) lower compared to the CLOs in the control group, while the average rating factor of their purchases is 166.4 (i.e., 67.5 - 6.9 + 105.8) higher. Column (3) in Table A8 allows us to analyse how OC constraints affect the portfolio composition of each individual CLO by aggregating across all the trades executed between reporting periods. The variable of interest is now the change in the WARF caused by all CLO trades between reporting dates, namely: $\Delta WARF_{i,t} = \sum_j RF_{i,j,t} \times \frac{\text{Amt. Purchased}_{i,j,t}}{\sum_j \text{Amt. Purchased}_{i,j,t}} - \sum_j RF_{i,j,t} \times \frac{\text{Amt. Sold}_{i,j,t}}{\sum_j \text{Amt. Sold}_{i,j,t}}$. As the previously results suggested, CLOs - when unconstrained - make their portfolios safer with their trades, as shown by the fact that, on average, they experience a change in WARF of 441.3 points. CLOs whose loans have been downgraded to Caa (CCC), however, deteriorate the quality of their assets by 316.2 point, while those whose Junior slack is between 0% and 5% by 374.5. CLOs whose constraints are binding and whose loans have been downgraded to Caa (CCC) do even worse by decreasing the quality of their portfolio by a further 79.1 points, implying that their assets are, on average, 769.9 points riskier compared to CLOs in the control group and 328.6 points compared to the counterfactual where they did not execute any trade.

To conclude, the evidence in this section suggests that when hit by downgrades to Caa (CCC), constrained CLO managers sell high quality securities in order to pass their OC test requirements compared to managers that decided to fail their tests or could not do otherwise. These loans are sold for non fundamental reasons, allowing us to use downgrades to Caa (CCC) as an exogenous source of variation in the supply of loans held by the CLOs affected by the shock. Consider, for instance, the following thought experiment: there are two ex-ante identical loans, L_A and L_B, which are held by CLOs α and β. An otherwise unrelated loan, say L_C, is downgraded to Caa (CCC). Imagine, L_C is held by CLO α, not by β. Once α is hit by a downgrade, L_A is more likely to be sold to restore α's OC constraint; if the market for L_A is not liquid enough to absorb the excess supply, price might fall compared to L_B as long as arbitrageurs are unable to spot the mispricing and bring it back in line. The usual issue with this type of reasoning is that α's management team endogenously chooses to sell L_A, and this might happen for reasons that are unobservable to the econometrician. This is, however, less likely to represent an issue in this case: we have shown in this section that OC tests incentivize managers to sell their best performing loans, implying that, if present, selection is likely to go against finding any price impact, that is, other investors should be willing to buy high quality assets which are sold by distressed CLOs. If, on the other hand, we find that the price of L_A is depressed compared to L_B, we should conclude that this is due to the price impact of CLOs and to the fact that it might take some time for arbitrageurs to direct their capital towards the secondary market for leveraged loans (Duffie, 2010). The next section will analyse and test this hypothesis in greater detail.

1.5 Impact on Prices

In the previous section I have shown how the presence of binding OC constraints can distort the trading behaviour of CLO managers; in particular I have shown that whenever a loan is downgraded below Caa (CCC), managers affected by the downgrade engage in par building by selling higher rated securities, especially so if their junior OC tests are binding. This trading behaviour gives us an instrument to investigate the extent to which the prices of loans are effected by forced sales that are unrelated to fundamentals. I can therefore trace the impact of the forced sale on the loan's price controlling for ex-ante measurable characteristics to get an estimate of how the secondary market for leveraged loans absorbs supply shocks. The endogenous choice of selling a security might bias the results towards finding no effect, given that loans that are sold by distressed CLOs tend to be of higher quality compared to the ones sold by non-distressed ones. One might still be worried that selection happens through unobservable characteristics that cannot be measured by the econometrician. To alleviate this concern, I show in Section 1.5.1 that this is not the case even when we assess the extent of ex-post selection: conditioning on ex-ante observables, loans that are sold by distressed CLOs are not more likely to be downgraded, to default, nor they display worsened liquidity. We can therefore conclude that the results in this section provide a lower bound for the impact of CLOs on the secondary market for leveraged loans.

In order to measure the impact of CLOs, I compare the average discount of loans sold by distressed CLOs after unrelated downgrades to Caa (CCC) with the average discount of a group of control loans. The identification assumption is that, after having controlled for time varying loan characteristics, the average discount of the two groups of loans would be identical if not for the fact they have been sold by distressed CLOs in order to restore their OC tests. The effect should be largest among loans that receive more selling pressure: for this reason I proceed by constructing a dummy variable, $\text{Shocked}_{j,t}$, that is equal to one whenever a loan has been sold by an above median number of distressed CLOs, where a CLO is considered distressed if their loans have been downgraded to Caa (CCC) or lower and the Junior OC slack is between 0% and 5%. These two conditions have been proven to be related to non fundamental par-building in the previous section. I then construct a dummy variable $\text{Post}_{j,t}$ that is equal to one in the twelve months after a loan has been sold by an above median number of distressed CLOs. I then proceed by running the following regression:

$$\text{discount}_{j,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.6}$$

The matrix $X_{j,t}$ contains various controls that are supposed to partial out the effect of measurable characteristics and exclude selection on ex-ante observables: the regressions include fixed effects for the time-to-maturity (TTM), the rating, the industry and the interest rate of the loan, all interacted with year×month fixed effects, assuring that the discount of treated loans is compared with loans with identical characteristics trading in the same month. I also include the lagged average discount on loans by the same issuer to control for the average discount on loans issued by the same company[15,16]. Column (1) in Table A9 shows that shocked CLOs

[15] Table A31 includes issuer fixed effects instead of the lagged average discount per issuer. The key difference between these two specification lies in the fact that the latter control for the average discount on the issuer's loans before and after the treatment date and, hence, might be influenced by the treatment itself. The results are, however, similar to the ones in Table A9 suggesting there is no selection in this direction.

[16] Figure A10 in the Appendix shows the fraction of variation explained by each fixed effect in-

sell expensive loans to improve their OC tests: loans sold by shocked CLOs trade at 39.5bps premium compared to their control group, confirming the findings in the previous section. This difference, however, becomes insignificant once we add time varying fixed effects for time-to-maturity, rating, industry and interest rate in columns (2)-(5), implying that most of the unconditional differences in prices between the treatment and control loans is absorbed by ex-ante measurable characteristics; this points to the fact that fixed effects take care of most of the selection between treatment and control groups in the choice of which securities should be sold. When we look at the impact of sales by shocked CLOs on prices, we can notice that treated loans trade at a discount of 71.8bps compared to untreated loans. The result is still statistically significant and large when we add the other fixed effects in columns (2)-(5), with the coefficients ranging between 43.6 and 49.4bps.

I then proceed to analyse whether loans sold by distressed CLOs trade at a significant discount compared to other loans issued by the same company k by including time varying issuer fixed effects which are meant to control for any unobservable time varying firm characteristics. I report in Table A10 the results of the following regression:

$$\text{discount}_{j,k,t} = \alpha_{k,t} + \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,k,t} \qquad (1.7)$$

where $\alpha_{k,t}$ are year×month×issuer fixed effects. Loans sold by distressed CLOs trade at 12bps discount compared to otherwise similar loans issued by the same borrower. Two considerations are however in order. First, the median (average) number of loans per issuer actively traded in any given month is 2 (3.1) and 33.6% of the time an issuer has a unique loan traded in a given month, implying that a large fraction of the observations do not have a loan in the control group. If we take into account the fact that companies with fewer traded loans tend to be smaller and less liquid, this might explain the reduction in magnitude of the coefficients compared to Table A9. Columns (2)-(4) in Table A10 confirm this hypothesis: when I restrict the sample to issuers having at least 2,5 and 10 actively traded loans

cluded in the regressions, ranging from year×month rate fixed effects, which explain 15% of the variation, to loan issuer × year × month fixed effects, which explain 90% of the variation.

the impact of CLOs grows to 12.6bps, 15.3bps and 17.9bps, respectively. Second, loans tend to be priced by comparables (Murfin and Pratt, 2018) and it is likely that the shock to a specific issue reverberates across all the loans issued by the same company even when this shock is non fundamental, proof being the fact that the issuer×year×month fixed effects explain a large fraction of the variation in discounts. The spillover from loans sold by shocked CLOs to loans issued by the same company implies that the estimate of β_2 is likely to underestimate the true magnitude of the effect in this specification. On top of that, notice that the $\text{Post}_{j,t}$ dummy takes into account the average difference in discounts in the twelve months following the shock: if the impact of the shocks lasts less than one year, we should expect this coefficient to be downward biased, given that it is already incorporating part of the reversal. In order to check whether this is the case, and trace the impact of a shock across time, I construct a set of dummy variables tracking the discount every month around the forced sale, which allows me to run the following regression:

$$\text{discount}_{j,t} = \gamma \text{Shocked}_{j,t} + \sum_{s=-6}^{12} \beta_s \text{Shocked}_{j,t} \times \mathbb{1}(t+s) + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.8}$$

where $\text{Shocked}_{j,t}$ is defined as above and $\mathbb{1}(t+s)$ is a set of dummies that are equal to one $s = -6, -5, ..., 11, 12$ months around the event of the sale at time t. The results are provided in Table A12 and plotted in Figure A5[17]. The results in Table A12 show that the direct impact of CLOs is between 22bps and 25.9bps in the month when loans are sold. The gap between the treatment and control group starts widening in the months after the sale with the difference reaching its maximum around six months after the sale when treated loans trade at between 54.4bps and 103bps discount compared to loans in the control group. Shocked loans need between ten and twelve months before their discount is not statistically different from the discount of the control group, but their difference starts to reduce significantly between the seventh and the ninth month. Figure A22 plots the average discount of loans in the treated group showing that their discount jumps imme-

[17] Figure A5 plots the results for model (4) in Table A12, including year×month×time-to-maturity, year×month×rating and year×month×industry fixed effects. The other results are plotted in Figure A21.

diately on impact; however, CLOs are more likely to have binding OC constraints in months when the market for leveraged loans is in distress. In the months after the shock, treated loans recover at a slower pace compared to those in the control group, widening the gap in average discounts up to seven/nine months when the treated loans slowly converge back to the control group. The presence of significant price reversal towards the control group is suggestive of the fact that the increase in discounts is more likely due to price pressure rather than changes in the fundamental value of the assets (Coval and Stafford, 2007); the fact that the market for leveraged loans is dislocated between seven and nine months after the fire sale is consistent with the evidence in Elkamhi and Nozawa (2021) who find a price impact of 35-40 weeks for loans sold by CLOs.

Finally, I turn to loans that distressed CLOs have purchased to build par. We can expect an asymmetry between sales and purchases in times of distress: distressed CLOs have a hard time finding investors willing to buy their assets, while it might be easier to meet with an investor willing to sell in periods of distress. I proceed to analyse the effect on loans purchased by distressed CLOs by running a difference-in-differences regression similar to the one in equation (1.5). If we look at the coefficient on $\text{Shocked}_{j,t} \times \text{Post}_{j,t}$ in Table A11 we quickly realize there is no evidence of upward price pressure on loans purchased by distressed CLOs. The coefficient is never statistically significant, with the exception of column (2) which includes only year$\times$month$\times$time-to-maturity fixed effects, suggesting that the discount on loans purchased by shocked CLOs is not different from other loans purchased in the twelve months after the event. The dynamic responses of discounts to forced purchases by shocked CLOs are reported in Table A33 and Figure A23.

Overall, the evidence in this section is conclusive of the fact that CLOs exercise a statistically and economically significant price pressure on the loans they are forced to sell whenever their OC tests bind because of a shock to their Caa (CCC) bucket. I have previously argued that any selection in the treated loans is likely to generate downward bias in the magnitude of the price pressure. The addition of fixed effects based on loan characteristics has made sure that, conditional on ex-ante observables, there is indeed no selection between treated and control groups. However, the following section tests whether there is any selection ex-post by analysing whether shocked loans are more likely to default or be downgraded.

1.5.1 Absence of Selection

By including time, time-to-maturity, rating, industry and borrower fixed effects we can be confident that the effect on loan discounts documented in the previous section is not motivated by ex-ante observable characteristics. In the specification where with year×month×issuer fixed effect, we can be confident that all the borrower time-varying characteristics have been partialled out in the analysis, hence showing that at least a fraction of the effect is security specific. However, we can still argue that CLOs might trade on ex-ante unobservable information which is not captured by the previously used controls. There is indeed evidence that trades in the secondary market for leveraged loans predict equity returns (Addoum and Murfin, 2020), suggesting that active investors in this market have access to private and valuable information; moreover Fabozzi et al. (2020) and Cordell et al. (2020) show that CLO managers tend to profit by actively trading loans, further suggesting that they might have access to information that is not available to other investors in real time. In order to differentiate between the hypothesis suggested in Section 1.5, i.e., that trades by distressed CLOs cause price pressure, and an alternative hypothesis where distressed CLOs simply have access to superior information and forecast a drop in the price of the loans they sell, we need to look at whether their trades are indeed able to predict outcomes that can be measured ex-post by the econometrician. In particular, in the rest of this section, I show that trades by distressed CLOs are not able to predict defaults or rating downgrades, therefore suggesting that these trades are indeed purely caused by concerns related to their OC tests.

The first test I conduct regards defaults. The results in Table A12 and Figure A5 suggest that the price impact of CLO trades lasts up to twelve months, therefore I proceed by constructing a dummy variable that is equal to one whenever a loan defaults in the following twelve months and check if loans sold by distressed CLOs are more likely to default compared to other loans held by CLOs. In order to control for loan characteristics by employing fixed effects and avoid the *incidental parameter problem* (Neyman and Scott, 1948), I will use the following linear probability model:

$$\text{default}_{j,t\to t+12} = \beta_1 \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.9}$$

where $\text{default}_{j,t\to t+12}$ is a dummy variable equal to one when loan j defaults in the period between t and $t+12$, $\text{Shocked}_{j,t}$ is a dummy equal to one when the loan has been sold by distressed CLOs and $X_{j,t}$ is a matrix of fixed effects to control for loan characteristics. The results are reported in Table A13. The baseline average probability of defaulting in the following twelve months for a loan that is sold by a CLO is equal to 4.79%; column (1) shows that this probability is reduced by 2.3% for the loans that are sold by distressed CLOs, confirming the hypothesis that shocked CLOs tend to sell higher quality securities to meet their OC test constraints. The effect is large compared to the baseline and significant even when we include time-to-maturity fixed effects at 0.9%. After the addition of rating fixed effects in columns (3)-(5) the coefficient becomes insignificant, suggesting that ratings capture the difference in risk between treatment and control group. Overall, the evidence in Table A13 is consistent with the hypothesis that sales by shocked CLOs have no informational motives and are executed to meet OC constrains, adding to the evidence that shocked CLOs are more aggressive in building par, they sell loans with lower rating factors, and they cause price pressure on the loans they sell.

I can further develop this hypothesis by testing whether loans sold by distressed CLOs are more or less likely to be downgraded or upgraded in the twelve months following the sale. In order to do so, I run the following regression:

$$\text{downgrade}_{j,t\to t+12} = \beta_1 \text{Shocked}_{jt} + X_{jt}\delta + \varepsilon_{j,t} \tag{1.10}$$

where $\text{downgrade}_{j,t\to t+12}$ is a dummy variable equal to one when loan j is downgraded in the period between t and $t+12$. Similarly I run a regression where the outcome variable is $\text{upgrade}_{j,t\to t+12}$, a dummy variable equal to one when a loan is upgraded in the period between t and $t + 12$. The results of the regressions are reported in Table A14. The three leftmost columns show that loans sold by distressed CLOs after downgrades to Caa (CCC) of otherwise unrelated loans are not more likely to be downgraded compared to loans in the control group after we control for ex-ante observable loan characteristics: in none of the specifications the coefficient is statistically different from zero. The same conclusion can be drawn by looking at upgrades: none of the specifications makes us conclude that treated loans are less likely to be upgraded in the following twelve months, compared to their control group. Ex-ante observable characteristics predict ex-post observable outcomes, suggesting that any regression including the former is unlikely to suffer from selection issues between treatment and control groups.

Finally, one might wonder whether loans sold by distressed CLOs differ from loans sold by other CLOs in terms of their liquidity. If loans sold by CLOs under duress become less liquid after the sale, and if liquidity commands a premium, the difference in discounts documented in Section 1.5 might be simply due to compensation for this type risk. In order to test this hypothesis, I will employ two commonly used proxies for liquidity in the literature on corporate bonds. First I use Roll (1984)'s measure computed as the the negative of the autocovariance in price changes, i.e. $\gamma = -Cov(\Delta p_t, \Delta p_{t-1})$. Second I use the number of trades per loan per month. The average value of γ is equal to 0.487, in line with measures for the corporate bond market (Bao et al., 2011), while the average number of trades per month is equal to 8.98, but heavily skewed towards few liquid loans, with the median number equal to 5 and way lower than the average number of trades in the corporate bond market. Table A15 reports the results of the following regression:

$$\text{liquidity}_{j,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.11}$$

where $\text{liquidity}_{j,t}$ is either γ or $\log(\text{Nr.Trades})$. In general we can conclude that there is no clear difference in liquidity after a loan has been sold by a distressed CLO. When we look at γ, the difference is significant in none of the specifications; moreover, a higher value of γ signals higher illiquidity, implying that - if something

- loans sold by shocked CLOs tend to be more liquid ex-post, even though this is not statistically significant. When we look at the number of trades, we cannot draw a firm conclusion: column (1) seems to suggest that loans sold by shocked CLOs trade 3.9% more often after the sale, however the coefficient switches sign and becomes insignificant in columns (2) and (3) once we add further controls, suggesting that there is no statistical difference in liquidity.

We can conclude this section by summarizing its main findings. While we have controlled for selection on ex-ante observables in Section 1.5, one might be worried that treatment and control groups could select on unobservables. In this section we have shown that, on the contrary, loans sold by shocked CLOs tend to be less likely to default after the sale, there is no statistical difference in the likelihood of being downgraded or in the liquidity of the loan ex-post. This confirms that selection on unobservable characteristics is unlikely to explain the findings in Section 1.5.

1.5.2 Placebo Tests

In order to make sure the results in Sections 1.4 and 1.5 are not spurious I proceed to conduct various placebo tests. The previous analysis is based on the premise that shocks to the Caa (CCC) bucket have potentially a material effect on OC tests and managers are forced to trade in order to make sure their tests are not violated. In order to test whether these distortions are really caused by the downgrades to Caa (CCC) I conduct the following placebo test: I construct a dummy variable that turns on when a CLO receives a shock to the bucket of securities rated B3 (B-) by Moody's (Standard & Poor's) and test whether these shocks have any effect on the trading behaviour of CLOs. Given that the OC tests are unaffected by the downgrade, the behaviour of management teams should not be distorted by these shocks. First I study whether CLOs whose loans have been downgraded to B3 build par by regressing the amount of par gained in each transaction on this dummy variable and another indicator that is equal to one whenever the slack of the CLO is between 0% and 5%. The results are presented in Table A16, from which it is clear that, as expected, CLOs hit by downgrades to B3 are neither more likely nor less likely to build par compared to CLOs in the control group: none of the coefficients is statistically significant and their magnitudes tend to be puny. This is indeed consistent with the idea that OC tests are insensitive to rating downgrades, as long as these downgrades do not affect Caa (CCC) buckets, implying that treated CLOs do not have any incentives to engage in par building more than they usually do. A similar placebo test where I consider as shocked those CLOs whose loans have been downgraded to a rating of B2 (B) according to Moody's (Standard & Poor's) is presented in Table A35, displaying similar results.

I then proceed with the final placebo test which is constructed similarly and tests whether sales performed by CLOs that have suffered a downgrade to B3 cause price pressure. As shown in Table A16, the trading bahaviour of these CLOs is not significantly different from the trading behaviour of CLOs in the control group, hence there is no reason to expect their trades will happen at depressed prices. Moreover, downgrades to B3 do not generate significant pressure across CLOs, implying that there will likely be other CLOs with similar portfolios willing to buy these loans. I construct a dummy variable, $\text{Shocked}_{j,t}$, that is equal to one if a loan has been sold by an above median number of distressed CLOs that have received downgrades to B3 and a dummy variable $\text{Post}_{j,t}$ that is equal to one after the loan has been sold by shocked CLOs. Table A17 reports the results of the following regression:

$$\text{discount}_{j,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times Post_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.12}$$

where the variables are constructed as in previous sections. The results in Table A17 confirm that loans sold by CLOs affected by downgrades to B3 do not trade at a significant discount compared to loans in the control group. A similar result for loans which have been sold by CLOs that suffered downgrades to B2 is reported in Table A36, with similar findings.

1.6 Impact on Primary Markets

Section 1.5 has documented that CLOs whose loans have been downgraded to Caa (CCC), and whose OC constraints are binding, sell securities for non fundamental reasons in order to restore the value of their tests. These sales depress the price of loans for up to seven/nine months. Section 1.5.1 has shown that the trading behaviour of distressed CLOs is not likely motivated by access to superior information. Finally, Section 1.5.2 has provided evidence that these results are specific to downgrades to Caa (CCC), hinting at the fact that these trades are indeed carried out in order to gain slack on OC tests. Overall, these results can be interpreted as evidence in favour of dislocations in the secondary market for loans unrelated to issuing companies' fundamentals and purely motivated by the mechanics of OC tests. However distortions in the secondary market might simply result in zero-sum transfers between distressed CLOs and unconstrained investors who are able to purchase high quality securities at depressed prices and, hence, have limited economic implications. That might not be the case whenever companies are forced to access dislocated markets and accept worse terms of financing compared to what they would have been able to do otherwise. Testing this hypothesis is particularly challenging: firms can endogenously reduce their demand for funds or divert it towards different markets, for instance, by trying to finance themselves using corporate bonds, equities or even other types of leveraged loans that are not affected by the previously documented shocks. In this section I will provide evidence in favour of the hypothesis that companies whose loans have been sold for non fundamental reasons face worse terms of financing in the leverage loans primary market. The demand for credit is an endogenous variable and if, as hinted above, companies whose loans have been sold by distressed CLOs are of higher quality, they might demand less capital in response. I attempt to tackle the issue of endogenous credit demand by adopting the following two strategies. First, I focus on companies whose previously issued loans are due to mature in the twelve months after they have been sold by distressed CLOs. As in Almeida et al. (2011), the fact that a firm is scheduled to refinance its debt in the following twelve months is the result of previous financing decisions and is, therefore, plausibly exogenous with respect to the downgrades to Caa (CCC) that have affected unrelated companies sharing the same CLO's balance sheet. By focusing on this subset of firms for the treatment and control group, I make sure that the two have roughly the same likelihood of requiring funds. Second, and more importantly, I look at firms that do eventually refinance themselves using leveraged loans and study the composition of newly issued securities between institutional and non-institutional loans. This strategy is reminiscent of the one used by papers analysing the substitution between corporate bonds and loans in periods of distress (Adrian et al., 2013; Becker and Ivashina, 2014) and by those on the substitution between institutional and bank tranches in the syndicated loan market (Ivashina and Sun, 2011a; Fleckenstein et al., 2020). This strategy guarantees that any variation in the terms of the loan cannot be explained by company-specific factors, given that institutional and bank tranches are claims on the same assets and usually have identical seniority[18].

I document four facts: first, companies whose loans have been sold by distressed CLOs face a higher cost of capital; second, these companies are less likely to issue institutional tranches which are usually held by CLOs; third, these companies borrow less money through institutional tranches; fourth, conditional on issuing an institutional tranche, these tend to be smaller. Figure A6 provides suggestive evidence in these regards, which I discuss in greater detail in the following paragraphs.

First, I study whether the price impact in the secondary market translates into higher cost of funding in primary markets. Panel (1) of Figure A6 plots the yearly average all-in spread drawn (AISD)[19] for companies that have been affected by fire sales from distressed CLOs (in red) against other companies (in blue); both groups are in the CLO-i/SDC Platinum matched sample, guaranteeing that, at least once, their loans have been held by CLOs. In each year in the sample the cost of capital for firms whose loans have been sold in the secondary market by distressed CLOs is higher by almost 100bps. This is not surprising given the evidence in Section 1.5: leveraged loans are usually priced by looking at the price of comparables (Murfin

[18]These, however, are different along other dimensions such as their pricing and, mainly, their amortization schedule.

[19]The all-in spread drawn is measured as the total annual spread including fees paid over the reference rate (usually LIBOR) for each dollar drawn from the loan. It therefore includes any fee or commission associated with the syndication process.

and Pratt, 2018) and the price of leveraged loans of the same company traded in the secondary market represents a clear benchmark for the cost of new issues (S&P Global, 2020b). Once previously issued loans trade in the secondary market at a discount, it is realistic to expect that new issued loans will likely be priced taking this extra spread into account. This is even less surprising once we recall that most of the discount suffered by these loans is due to a lack of convergence towards the control group: after an initial negative shock, the discount on treated loans reduces but in a slower fashion compared to their control group, making the mispricing particularly hard to detect to an investor that is simply looking at the treated loans in isolation. The results in Figure A6 do not control for firm characteristics, which might drive the difference in AISD between the two groups. For this reason, I investigate the effect on spreads by running the following regression whose results are reported in Table A19:

$$\text{AISD}_{j,t} = \alpha_t + \alpha_j + \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.13}$$

where $\text{AISD}_{j,t}$ is the all-in drawn spread for issuer j at time t and $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by distressed CLOs in the previous twelve months. Column (1) shows that the spread for shocked firms is 55bps higher compared to the control group once we include time fixed effects which help in partialling out any shared macroeconomic variation in loan spreads. This guarantees that the results do not stem from the fact that downgrades to Caa (CCC) are more likely to happen during bad periods, since we are focusing on the cross-sectional differences in AISDs. Columns (2)-(4) show that the effect is robust to the inclusion of time-to-maturity, industry, rating and issuer fixed effects even though the magnitude of the coefficients is reduced. The effect ranges between 23.2 and 34.7 basis points when we consider time-to-maturity, industry and rating fixed effects. When we add issuer fixed effects in column (5), the magnitude is reduced to 8bps, however it should be noted that the effect is now identified from the subset of firms that have issued loans twelve months after being sold by CLOs and have been at least once in the control and at least once in the treated group, representing a small subset of the whole sample. A significant fraction of the shock documented in Section 1.5 is transmitted to the primary market of issuance, implying that the ex-post cost of borrowing for treated firms is higher than for firms in the control group. These firms are facing a shock to the supply of institutional loans and might be forced to finance themselves with the next best source of capital. The rest of the section tests this hypothesis by looking at differences in the composition of borrowed funds between the two groups of companies. Panel (2) in Figure A6 compares the number of institutional tranches as a fraction of the total number of tranches issued by the two groups every year. With the exception of 2009 and 2010, the fraction of institutional tranches for shocked issuers has always been lower compared to firms in the control group, corroborating the hypothesis that companies indeed move away from institutional loans whenever they face a supply shock in this market. Notice that by focusing only on firms that do eventually issue leveraged loans we are able to fix the demand for funds and make sure that indeed the effect is driven by the supply of credit. Panel (3) provides similar evidence if we look at the total amount of funds borrowed using institutional tranches as a fraction of the total amount borrowed: after 2010 shocked firms have borrowed less capital compared to firms in the control group via institutional loans.

I then proceed to test whether the fraction of institutional tranches and the amount of dollars borrowed are significantly lower for treated firms by running the following regressions:

$$\text{Fraction Inst.}_{j,t} = \alpha_t + \alpha_j + \beta \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.14}$$

$$\text{Fraction Inst. \$}_{j,t} = \alpha_t + \alpha_j + \beta \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t} \tag{1.15}$$

where Fraction Inst.$_{j,t}$ measures the number of institutional tranches as a fraction of the total number of tranches issued by issuer j at time t and Fraction Inst. \$$_{j,t}$ measures the total amount of dollars borrowed using institutional tranches as a fraction of the total amount borrowed by issuer j at time t. The results reported in Table A20 and Table A21 show that shocked borrowers issue between 3.9% and 11.7% less institutional tranches, while the amount borrowed is between 5.5% and

11.3% lower after we have controlled for firm characteristics, suggesting that affected borrowers are substituting between institutional and traditional bank loans. How can we interpret this evidence? If companies finance themselves with the cheapest source of financing trying to arbitrage across different types of securities (Ma, 2019), when the supply of institutional loans shifts they might move to the next best source of funds, namely traditional (non institutional) bank loans. Once forced to access the next source of capital we should expect the cost of capital to increase, as documented in Table A19, and the quantities borrowed to decrease.

As a final test, Panel (4) in Figure A6 and Table A22 study intensive margin effects, by looking at the size of institutional tranches in terms of dollars borrowed: even when they maintain access to institutional loans, treated companies draw less funds through institutional loans sold to CLOs. The institutional tranches of companies affected by dislocations in the secondary market are between 23.5% and 34.2% smaller compared to those in the control group. This is true even when we include issuer fixed effect, which control for the average size of loans issued by that specific borrower.

1.7 Conclusions

The paper studies the effect of non fundamental trades executed by CLOs in order to gain slack in their constraints. After having analysed which loans are sold by distressed CLOs in order to restore their constraints, I study the impact of their trading actions. Securities sold by distressed CLOs trade at roughly 40bps discount: I show that this effect is likely causal and cannot be explained by selection on ex-ante or ex-post measurable loan characteristics. The effect is long lasting (up to nine months) and mostly due to the failure of loans sold by distressed CLOs to recover from depressed prices. I provide evidence that shocks in the secondary market transmit to the primary market: companies that are due to refinance their loans are less likely to employ institutional tranches when hit by the selling pressure of CLOs in the secondary market. The substitution away from institutional to bank tranches increases the cost of capital for affected firms.

On top of documenting the relative price inelasticity of the market for leveraged loans and the downstream effects on companies' financing decisions, the results in the paper might inform regulators when designing constraints for institutional investors: constraints based on the par value of assets may prevent investors from engaging in selling spirals where shocked assets are sold by distressed investors, further exacerbating the initial shock. However, they might also lead to spillovers where shocks are transmitted from troubled securities to otherwise unrelated ones through the balance sheet of institutional investors. Finding the optimal balance between the two concerns is crucial and should be the topic of further research.

2. Revealed Expectations and Learning Biases: Evidence from the Mutual Fund Industry

FRANCESCO NICOLAI AND SIMONA RISTESKA[1]

How do investors form their return expectations? Do they take all available information into account? Does personal experience play a crucial role in the formation of expectations? We attempt to answer these questions by looking at mutual fund managers' stock return expectations as revealed by their portfolio holdings. We exploit the fact that, under a large class of models, the optimal portfolio rule has a similar functional form; using a three dimensional panel consisting of the portfolio holdings of mutual fund managers over a period of thirty-five years, we are able to extract a measure of subjective expected returns for every manager in our panel by exploiting the variation across stocks over time between and within managers. To see this, consider a mean-variance investor for whom the vector of physical expected returns is given by the following formula:

$$\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}] = \gamma_{i,t}\Sigma_t\boldsymbol{w^*_{i,t}} \tag{2.1}$$

where $\mathbb{E}_{i,t}[\cdot]$ is the conditional expectation operator taken under investor i's information set at time t, $\boldsymbol{r_{t+1}} - r_f\mathbf{1}$ is a vector of excess returns, $\gamma_{i,t}$ is the coefficient of relative risk aversion of manager i at time t, Σ_t is the conditional covariance matrix of stock returns and $\boldsymbol{w^*_{i,t}}$ is the time t vector of optimal portfolio weights of investor i. The above expression for expected excess returns is obtained by invert-

[1]We benefited from helpful comments from Ulf Axelson, Nicholas Barberis, Pasquale Della Corte, Daniel Ferreira, Boyan Jovanovic, Christian Julliard, Samuli Knupfer, Ralph Koijen, Avner Langut, Dong Lou, Ian Martin, Igor Makarov, Cameron Peng, Asaf Razin, Andrew Redleaf, Andrea Tamoni, Michela Verardo and the participants at the LSE seminar, the 2019 Yale Whitebox Conference, the 2019 Belgrade Young Economists Conference. Any errors or omissions are the responsibility of the authors.

ing the first-order condition of a mean-variance investor provided that we have a good measure of conditional covariances Σ_t. In these regards we follow Merton (1980) and argue that investors' disagreement should mainly regard expected returns and not variances and covariances. We show that empirically this is a good approximation. In Section 2.2.1 we show that - as long as we correctly interpret the manager-specific time-varying parameter $\gamma_{i,t}$ - many optimal portfolio models give rise to a subjective expected return similar to (2.1); whenever that is not the case, we can saturate the model with fixed effects in order to split the total demand into a mean-variance component and a hedging component; to isolate the effect of risk aversion from the effect of subjective expected returns we resort to the very general principle that, given the cross-section of assets the manager invests in, risk aversion is a manager-specific quantity, while expected returns are at the same time asset-specific. The information contained in the cross section of holdings, therefore, greatly reduces the issue of separating the variation due to the manager's preferences from the one due to beliefs.

We start by providing evidence in Section 2.3 that more than 50% of the variation in expected returns is explained by a common time-varying factor and we are able to explain almost 90% of the variation with manager-time and stock-time components. This suggests that a saturated regression will likely allow us to isolate the idiosyncratic part of expected returns affected by manager-stock-time specific effects; we focus on this component to explore the extent to which managers' beliefs are affected by experience. In particular, we investigate whether fund managers put more emphasis on past stock returns that they have *personally* experienced over their investment career. To begin, we consider the effect of the simple average of past observed returns on portfolio holdings decisions. Having experienced a one standard deviation higher average return on a given stock causes the manager to inflate his expected excess return by between 10.3 and 15.1 basis points (after partialling out the effect of common stock and manager characteristics). This effect is both statistically and economically significant and it is almost an order of magnitude larger than that of other commonly used predictors. Nonetheless, the effect of average experienced returns masks important heterogeneity in the influence of past returns observed at different points in time: when we move on to examining the particular shape of the learning curve we find evidence of a differential impact. We start by providing non-parametric results that do not require taking a stance on the precise functional form that investors use to weight past experienced returns. Mutual fund managers in our sample are subject to the so-called serial-position effect: the tendency to predominantly remember the initial and the last observations in a series. More precisely, managers' investment decisions and beliefs are particularly affected by the returns they have experienced early on during their stock-specific experience and those they have experienced most recently. In other words, professional investors seem to exhibit the *primacy* and *recency* bias.

As one would expect, the effect is stronger for single-managed funds and decays fast as the number of managers in a team increases: the effect of recently experienced returns on managers in a single managed fund is twice as large compared to managers working with at least one other professional; the effect of early returns is an order of magnitude larger.

We also show that the differential effect of taxes on capital gains and losses cannot explain these findings since the effect of early career experience is still present even when the manager switches to a different fund. At most, tax considerations can explain 20% of the estimated influence of past returns on portfolio choices and expected returns.

Armed with the reduced-form evidence, we provide a tentative parametric estimation of the managers' learning function. In particular, the results in the reduced-form estimation seem to suggest a non-monotonic learning function. For this reason we adopt a variation of the parametrisation of the learning function in Malmendier and Nagel (2016) that allows for a variety of decreasing and increasing, convex and concave, monotone and non-monotone learning weights. We find that fund managers on average do indeed place a disproportionate weight on personal past experience and that this biases the expected returns recovered from their stock holdings, after having adjusted for risk and risk aversion. When we allow for time-varying weights on past stock returns, we show that mutual fund managers tend to place excessive weight on returns experienced at the beginning of their careers and in the most recent quarters compared to those in the middle period, suggesting that both early-career and recent experience seem to be important determinants of the investment behaviour of a large class of professional investors. For instance, a manager with the median stock-specific experience of 9 quarters assigns around 1.84 times larger weight on the return experienced in the most recent quarter compared to the benchmark of 1/9, while the weight on the first experienced return is 3.13 times larger than the benchmark. We thus reconcile two conflicting strands of the literature: similarly to Malmendier and Nagel (2011) and Malmendier and Nagel (2016), we confirm that investors do overweight their personal experience and manifest a *recency bias*, but - at the same time - we show that professional investors also place a disproportionately large weight on returns that have been experienced in the early part of their investing career, similarly to the findings of Kaustia and Knüpfer (2008) and Hirshleifer et al. (2021). When looking at co-managed funds, we show that a large fraction of the effect of early experience washes out while the effect of recently experienced returns persists; this might be due to the fact that, while there is large heterogeneity in early experience, recently experienced returns are mostly shared among managers within a team.

Finally, in the last part of the paper we focus on risk preferences. Notice from equation (2.1) that, while risk aversion varies at the manager-time level, beliefs vary at the manager-time-stock level. This lets us separate *variation* in adjusted portfolio holdings that is due to the managers' risk appetite from differences in beliefs, but does not inform us regarding their *level*. Once we make some minimal assumptions to pin down their level, we show that individual expected returns tend to be quite biased and that preferences display significant heterogeneity across individuals and time. Moreover, on average, mutual fund managers display an Arrow (1965)-Pratt (1964) coefficient of relative risk aversion between 0.915 and 1.283.

The rest of the paper is organised as follows: Section 2.1 provides an overview of recent literature. We proceed by showing that most of the literature relies on evidence from surveys obtained from non-professional investors or, when not affected by these concerns, on a relatively limited amount of data. We argue that the present paper tries to solve the aforementioned issues. Section 2.2 describes how we can separate the variation in expected returns from the variation in risk aversion or other factors in a wide class of models. Section 2.3 gives details of the data used in our empirical work and provides some summary statistics. Section 2.4 provides the non-parametric results of our analysis, while Section 2.5 describes and show the results of our parametric approach. In Section 2.6, we tackle the question of the level of risk aversion of investment professionals. Finally, Section 2.7 provides concluding remarks.

2.1 Previous Literature

The issue of whether economic agents learn with experience has been explored to some extent by the existing literature. Evidence from the literature in psychology and economics shows that personal experience exerts a larger influence on behaviour compared to other shared sources of information[2], especially very recent and very early experience. These two phenomena are usually referred to as the *recency* and the *primacy* effect and they generate what is known to researchers in psychology as the U-shaped serial-position curve[3].

Diving deeper into the field of finance there is growing evidence that personal experience affects financial behaviour. Kaustia and Knüpfer (2008) and Chiang et al. (2011) show that the likelihood of participating in subsequent IPOs is affected by returns experienced in previous offerings. Choi et al. (2009) provide evidence that investors with high return or low volatility on their 401(k) savings tend to invest a larger fraction of their wealth. Using data from the Survey of Consumer Finances from 1960 to 2007, Malmendier and Nagel (2011) find that individuals

[2]For early evidence on the concept of reinforcement learning, see the seminal study by Thorndike (1898). A large body of theoretical and empirical literature studies the role of personal experience in learning, see, for instance, Tversky and Kahneman (1973) for a discussion of the availability bias, Fazio et al. (1978) for experimental evidence on the differential processing of information that results from direct versus indirect experience, Roth and Erev (1995) and Erev and Roth (1998) for experimental data and theory regarding learning in sequential games, Camerer and Ho (1999) for a combined model of reinforcement and belief-based learning, Simonsohn et al. (2008) for experimental analysis of the effect of personal experience in a game theory context.

[3]The psychology literature on these topics goes beyond the scope of this paper. Among others, see Nipher (1878), Ebbinghaus (1913) and Murdock (1962) for evidence on the serial-position effect; for evidence on the primacy effect, see Asch (1946); the recency effect is explored by Deese and Kaufman (1957). See Murdock (1974) for a survey.

born before the 1920s who have experienced the lackluster stock market returns during the Great Depression report higher risk aversion, lower expected returns and are less likely to invest in the stock market. Those that happened to experience lower bond market returns tend to reduce their bond holdings. They also find that returns experienced in the previous year contribute four to six times more to future investment decisions than those experienced thirty years ago. In a similar vein, Malmendier and Nagel (2016) analyse the effect of life-time experience on inflation expectations using the Reuters/Michigan Survey of Consumers; they show that the effect is stronger for younger respondents, and has a direct effect on their borrowing and savings decisions. Malmendier et al. (2021) analyse the effect of experienced inflation on members of the FOMC board and find similar results. Greenwood and Nagel (2009) investigate the effect of experience on mutual fund managers during the dot-com bubble of the late 1990s. The authors use age as a proxy for experience and show that younger managers were investing more in technology stocks compared to similar older managers and displayed a more pronounced trend-chasing behavior. Chernenko et al. (2016) study the effect of experience on a panel of mutual funds holdings of MBS during the 2003-2007 mortgage boom and show that less experienced managers had larger positions in these securities, especially those backed by subprime mortgages; moreover they show that personal experience outside of the fund had an effect on portfolio choice behaviour. Andonov and Rauh (2020) analyse the effect of experienced returns on a cross-section of U.S. Pension Fund managers, showing a significant effect of past experience on the expected returns that these investors report in annual target asset allocations; in particular, earlier experiences have a stronger effect on investment behaviour. Giglio et al. (2021) look at retail investors' portfolio allocations and match them to beliefs elicited from surveys. They find that stated beliefs have a low explanatory power for the timing of trades, however, they are able to predict the direction and size of those trades that do occur. Finally, there is evidence that experienced risk affects financial behaviour: Knüpfer et al. (2017) show that experienced labour market distress affects portfolio choices, while Lochstoer and Muir (forthcoming) find that individuals have extrapolative beliefs about market volatility.

While the contribution of the above papers is substantial, we argue that most of them are affected by one or more of the following issues: reliance on evidence obtained from surveys where agents report their subjective expected returns, focus on non-professional investors who spend limited time investing and, usually, invest relatively small amounts, and reliance on limited time-series or cross-sections implying that it is harder to perform statistical inference.

Regarding the first issue, the task of recovering investors' expectations is a particularly tricky one. It is well known at least since Harrison and Kreps (1979) that asset prices reveal only risk-neutral expectations of market participants; a way to circumvent this problem is, therefore, to focus attention on expectations elicited from surveys. Most of these measures seem to display high correlations as Greenwood and Shleifer (2014) point out. However Cochrane (2017) argues that there is no guarantee that people report their "true-measure unconditional mean" in surveys. In these regards, Adam et al. (2021) provide evidence that surveyed expected returns are inconsistent with risk-neutral expected returns, ambiguity averse/robust expected returns or any other risk-adjusted returns[4]. However, nothing guarantees that the reported expected returns are exactly representative of the mathematical physical expectation of investors. Consider for instance a survey respondent that interprets the question as asking "what is the most likely return" instead of "what is the expected return". In that case, the respondent will provide a measure of the modal return rather than its average taken across states of the world. Although the previous example may seem far-fetched, Martin (2017) shows that - for a log investor who holds the market - the physical distribution of returns is asymmetric and, for instance, at the height of the crisis, while the expected return on the S&P 500 was above 20% per year, the author recovers a probability of almost 20% of a 20% decline in the index. Large probability masses far from the mean imply large discrepancies between modal, median and average returns. Beliefs reflected in portfolio choices are more informative and represent the primary object of interest, given that it is ultimately changes in demand and supply

[4] Appendix B.3 shows that our framework can also deal with this type of preferences.

that determine the variation in prices. Malmendier and Nagel (2011), Andonov and Rauh (2020) and Giglio et al. (2021) show that portfolio choices are consistent with stated beliefs, but the explanatory power is only partial, while - by construction - our beliefs are fully consistent with trading behaviour.

Regarding the second issue, we argue that there are reasons to believe that sophisticated professional investors might behave differently compared to households and, for this reason, we focus our attention on mutual fund managers; they also routinely follow the stock market and therefore there might be reasons to expect them to be less prone to biases or memory issues. While this seems to be true in the case of IPO subscriptions (Chiang et al., 2011), we show that our investors display large biases even though we cannot provide a direct comparison to households. It should also be noted that, to the extent that these financial intermediaries represent a large fraction of total stock market activity, their beliefs will be an important driver of stock price movements.

Finally, concerning the third issue, many of the papers dealing with institutional investors focus on specific events (e.g., Greenwood and Nagel (2009) or Chernenko et al. (2016)) or rely on limited time series data (e.g., Andonov and Rauh (2020)). The aim of the present paper is to be more general and explore whether the effect of experienced returns is common across periods and stocks and represents a permanent trait of professional investors' behaviour.

2.2 Methodology

In this section we provide a detailed description of our empirical strategy. We first explain how we obtain a measure of expected returns given portfolio holdings. We argue that in a wide set of models - including a mean-variance benchmark - we are able to separate the effect of risk and risk aversion from the effect of return beliefs by using the cross-section of manager holdings. We then describe the way we deal with the issue of estimating covariance matrices and, finally, our plan for identifying risk aversion.

2.2.1 Recovering Subjective Expected Returns

Portfolio choices reveal information about future stock return expectations: this is the main insight of Sharpe (1974)'s *indirect approach* to mean-variance optimisation whereby beliefs about expected returns are inferred from portfolio holdings, rather than the other way around[5]. Consider the problem of investor i who maximises his value function by choosing his portfolio allocations into a risk-free and N risky assets:

$$\max_{\{\boldsymbol{w_{i,t}},\dots\}} J_{i,t}(W_{i,t}) \tag{2.2}$$

where $J_{i,t}(.)$ is the value function of the investor evaluated at his current wealth $W_{i,t}$. When returns follow a geometric Brownian motion, the law of motion for wealth is:

$$\frac{dW_{i,t}}{W_{i,t}} = r_f dt + \boldsymbol{w}'_{\boldsymbol{i,t}}(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})dt - \Delta C_{i,t}dt + \boldsymbol{w}'_{\boldsymbol{i,t}}\Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{2.3}$$

where r_f is the instantaneous risk-free rate (or the instantaneous rate of return of any other *reference* asset with respect to which excess returns are computed), $\boldsymbol{\mu_{i,t}}$ is an $N \times 1$ vector of stock return drifts as perceived by investor i, $\boldsymbol{w_{i,t}}$ is an $N \times 1$ vector of stock portfolio weights, $\Sigma_t^{\frac{1}{2}}$ is an $N \times N$ matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, $\Delta C_{i,t}$ is the (net) outflow of resources[6], and $\mathbf{1}$ is an $N \times 1$ vector of ones.

The investor chooses his optimal portfolio by selecting $\boldsymbol{w_{i,t}}$. Notice that we are deliberately vague about other potential choice variables, i.e., our analysis follows solely from the optimality conditions for the portfolio holdings and the fact that current wealth is the only state variable. Standard dynamic optimisation argu-

[5] Black and Litterman (1992) start from the same insight to obtain portfolio holdings that combine the manager's views with average realised returns in a consistent way; Cohen et al. (2008) and Shumway et al. (2011) use a similar approach to extract a measure of beliefs from portfolios holdings. The former paper measures the *best ideas* of mutual funds as the investment positions for which the authors can extract the largest expected returns, while the latter analyses the rationality implications of extracted beliefs.

[6] For a standard consumption maximisation problem we can interpret $\Delta C_{i,t} = \frac{C_{i,t} - Y_{i,t}}{W_{i,t}}$, i.e., the instantaneous flow of consumption $C_{i,t}$ net of the income flow $Y_{i,t}$, expressed as a fraction of wealth $W_{i,t}$. In this setting $\Delta C_{i,t}$ can be loosely interpreted as the net outflow of money the mutual fund manager is subject to in each period because of redemption/creation of new fund shares. Because of Markovianity we have that $\Delta C_{i,t} = \Delta C(W_{i,t})$.

ments (Back, 2017) give the following optimality condition:

$$\boldsymbol{w}_{\boldsymbol{i,t}}^{*} = -\frac{J_{W_{i,t}}}{W_{i,t} J_{W_{i,t} W_{i,t}}} \Sigma_t^{-1} (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1}) \tag{2.4}$$

where $J_{W_{i,t}}$ and $J_{W_{i,t}W_{i,t}}$ are the first and second derivatives of the value function with respect to current wealth and therefore $-\frac{W_{i,t} J_{W_{i,t} W_{i,t}}}{J_{W_{i,t}}}$ is the Arrow (1965)-Pratt (1964) coefficient of instantaneous relative risk aversion measuring the curvature of the value function with respect to wealth, which we denote $\gamma_{i,t} \equiv -\frac{W_{i,t} J_{W_{i,t} W_{i,t}}}{J_{W_{i,t}}}$. Notice that equation (2.4) is a generalisation of the optimal demand employed by Koijen and Yogo (2019)[7]. We can invert the optimality condition (2.4) in order to get an expression for expected excess returns as a function of optimal holdings and Σ_t. In particular, we have that:

$$\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1} = \gamma_{i,t} \Sigma_t \boldsymbol{w}_{\boldsymbol{i,t}}^{*} \tag{2.5}$$

If we had information about the level of the investor's risk aversion $\gamma_{i,t}$ and the covariance matrix Σ_t, we could obtain an exact measure of his subjective expectations of future one-period ahead excess returns $\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1}$. We follow Merton (1980) in arguing that investors should share beliefs regarding Σ_t; we later provide evidence in support of this assumption. To isolate the effect of $\gamma_{i,t}$, let us consider each element of the vector of excess returns $\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1}$. At each point in time t, for each stock j, each manager i forms a measure of expected excess return which we can denote by $(\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1})_j$[8]. By simply keeping track of the subscripts one can realise that there is variation in expected returns across managers, stocks and time, i.e., along the three dimensions i, j, t. On the other hand, the coefficient of relative risk aversion $\gamma_{i,t}$ varies only at the i-t level, implying that the cross-section of holdings for manager i at time t gives us enough information to isolate the variation in beliefs from the variation in risk aversion which acts as a level shifter on the demand

[7] The optimal demand in equation (7) of Koijen and Yogo (2019) is equivalent to our specification whenever $-\frac{W_{i,t} J_{W_{i,t} W_{i,t}}}{J_{W_{i,t}}} = 1$, i.e., investors have logarithmic utility. It is easy to incorporate short sale constraints in our setting as we show in Appendix B.3.

[8] $(\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1})_j$ is the $j-$th element of the vector of expected excess returns for manager i, time t, i.e., $\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1} = [(\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1})_1, ..., (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1})_j, ..., (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f \mathbf{1})_N]'$.

for risky assets[9]. When instantaneous returns are normally distributed and wealth is the only state variable, any utility function (and therefore any value function $J_{i,t}(W_{i,t})$) gives rise to a demand as the one in (2.4). We can extend this approach to a wide class of models where there is an $L \times 1$ vector of Markovian state variables $\boldsymbol{X_t}$ with the following law of motion:

$$\boldsymbol{dX_t} = \phi(\boldsymbol{X_t})dt + \Gamma(\boldsymbol{X_t})\boldsymbol{dZ_t} \tag{2.6}$$

Standard dynamic optimisation arguments imply that, in that case, the optimal demand will be:

$$\boldsymbol{w^*_{i,t}} = -\frac{J_{W_{i,t}}}{W_{i,t}J_{W_{i,t}W_{i,t}}}\Sigma_t^{-1}\left((\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}) - \sum_{l=1}^{L}\frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}}\boldsymbol{K_{l,t}}\right) \tag{2.7}$$

where $\frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}} = \frac{\partial \log J_{W_{i,t}}}{\partial X_{l,t}}$ measures the semi-elasticity of the marginal utility of wealth $J_{W_{i,t}}$ with respect to the Markovian state variable $X_{l,t}$, and $\boldsymbol{K_{l,t}} = \Sigma_t^{\frac{1}{2}}\boldsymbol{\Gamma_{l,t}}$ represents the vector of instantaneous covariances between returns and the state variable $X_{l,t}$. Let us denote the hedging demand $\boldsymbol{H_{i,t}} \equiv \sum_{l=1}^{L}\frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}}\boldsymbol{K_{l,t}}$. There are many settings in which we can still disentangle variation in beliefs from variation in hedging demands[10]. First, we might consider the possibility that the mutual fund is facing borrowing constraints. We show in the Appendix that in this case the expected return can be recovered from:

$$(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_j = \gamma_{i,t}\left(\Sigma_t\boldsymbol{w^*_{i,t}}\right)_j + H_{i,t} \tag{2.8}$$

Similarly, suppose mutual funds managers are ranked according to a common summary statistic (e.g. alpha over a benchmark). The expected excess return can then be approximated by:

$$(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_j = \gamma_{i,t}\left(\Sigma_t\boldsymbol{w^*_{i,t}}\right)_j + H_{j,t} \tag{2.9}$$

[9]For the reader who is familiar with the textbook mean-variance optimisation, this is analogous to the fact that the selection of the tangency portfolio does not depend on the investor's risk aversion which merely influences the relative proportion of wealth invested in the risk-free and risky assets.

[10]For more details, see Appendix B.3 where we analyse the case of borrowing and short selling constraints, concerns about model misspecification and the issue of benchmarking.

The previous examples show that, by saturating the regressions with the proper fixed effects, we are able to use the cross-section of assets of a particular investor to separate the effect of changes in beliefs (which vary at the i, j, t level) from the effect of changes in risk aversion (varying at the i, t level) and hedging demand (as long as this varies at a coarser level).

As a caveat, notice that the only situation where we would be unable to separate changes in the hedging demand from changes in beliefs is if the hedging demand varied at the stock-manager-time level (i.e., if we had $H_{i,j,t}$). This would undermine any attempt to recover variation in beliefs from variation in portfolio holdings; however, the results in the paper would not lose their relevance. First of all, as shown by Moreira and Muir (2019) in the case of time-varying expected returns and volatilities, optimal portfolios can be closely approximated by an affine transformation of the standard mean-variance portfolio. Second, even if expected excess returns cannot be separated from hedging demands, it is not easy to conceive of a mechanism where past experience has a large impact on hedging demands. Third, even if this were the case, we could still interpret all the results in terms of scaled demands ($\Sigma_t \boldsymbol{w}_{i,t}^*$) as opposed to beliefs. Asset prices are ultimately determined by investors' holdings and the variation thereof; it would be nice to know whether the effect on investors' demands goes through expected returns ($\boldsymbol{\mu}_{i,t} - r_f \mathbf{1}$), risk aversion ($\gamma_{i,t}$) or hedging demands ($\boldsymbol{H}$), but ultimately what matters is the fact that part of the variation in the cross-section and the time-series of assets holdings is due to the returns that the agent has experienced. Having said that, in what follows, we impose the previously discussed restrictions in order to disentangle the different mechanisms. We, therefore, assume that the issue of hedging demands can be solved by saturating the regression with the appropriate levels of fixed effects. In the following two sections, we tackle the two remaining problems, namely, the estimation of the conditional covariance matrix and level of risk aversion.

2.2.2 Estimating the Covariance Matrix

As can be seen in the previous section, in order to construct a measure of one-period ahead expected excess returns, we need to have a measure of the conditional covariance matrices. In this paper we rely on an argument set forth by Merton (1980), which states that, in principle, all investors should agree on Σ_t since it can be very precisely estimated by using increasingly more granular data. In practice it is unavoidable to take a stance on how to estimate the conditional covariance matrix. To make sure that our results do not depend on the chosen estimator for Σ_t, we decide to take three different approaches for this exercise:

1. As a first measure, we compute the sample covariance matrix of stock returns:

$$\hat{\Sigma}_t^{d,1} = \frac{1}{t-1}(R_t - \bar{\boldsymbol{r}}_t \mathbf{1}')(R_t - \bar{\boldsymbol{r}}_t \mathbf{1}')'$$

   where $R_t = [\boldsymbol{r}_{1,t}, ..., \boldsymbol{r}_{j,t}, ..., \boldsymbol{r}_{N,t}]'$ is an $N \times t$ matrix that contains past realised returns as rows, $\bar{\boldsymbol{r}}_t$ is an $N \times 1$ vector that collects sample average returns computed at time t, and $\mathbf{1}$ is a $t \times 1$ vector of ones. We estimate $\hat{\Sigma}_t^{d,1}$ from a one-year rolling window of daily returns[11] and we scale it by $K = \frac{\text{nb. obs.}}{\text{nb. quarters}} = 63.07$ days to obtain our first estimator as $\hat{\Sigma}_t^1 = K \times \hat{\Sigma}_t^{d,1}$. It is well known that it is extremely hard to estimate correlations between stocks and correlations close to unity in absolute value tend to give extreme long-short portfolios. For this reason we resort to the next two measures of the sample covariance matrix;

2. Our second estimate makes use of a Bayesian Stein Shrinkage estimator. We follow Touloumis (2015) and compute the daily covariance matrix $\hat{\Sigma}_t^{d,2}$ as a weighted-average of the sample covariance matrix $\hat{\Sigma}_t^{d,1}$ and a target matrix

[11] The reader might be worried about the fact that we estimate expected returns employing covariance matrices that rely on past return data, to subsequently regress on past realised returns. However, notice that the same covariance estimates are shared in the cross-section of managers, which is not true for past experienced returns. Furthermore, our estimates of covariance matrices employ only one year of data while the average manager has more than three years of experience with a given stock.

Σ_t^{target} which imposes zero correlations across stocks:

$$\hat{\Sigma}_t^{d,2} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_t^{target}$$

where Σ_t^{target} is a diagonal matrix where the elements on the diagonal are the sample estimated variances, namely $\Sigma_t^{target} = \hat{\Sigma}_t^{d,1} * I_N$ where $*$ denotes the Hadamard product and I_N is a $N \times N$ identity matrix where N is the number of stocks. The estimator of quarterly covariances is then: $\hat{\Sigma}_t^2 = K \times \hat{\Sigma}_t^{d,2}$;

3. In our third and final approach, we again apply a similar Bayesian Stein Shrinkage Estimator:

$$\hat{\Sigma}_t^{d,3} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_t^{target}$$

Following Ledoit and Wolf (2004), $\tilde{\Sigma}_t^{target}$ is a diagonal matrix with the constant average daily sample variance on the diagonal, namely $\tilde{\Sigma}_t^{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the covariance matrix, and I_N is a $N \times N$ identity matrix where N is the number of stocks. The estimator is then: $\hat{\Sigma}_t^3 = K \times \hat{\Sigma}_t^{d,3}$.

More details on the construction of $\hat{\Sigma}_t^2$ and $\hat{\Sigma}_t^3$ and the optimal choice of λ are provided in Appendix B.4. We show in the rest of the paper that the way we compute the covariance matrices is not very relevant for our results. This should be expected given that, as long as managers' estimates of covariances are very similar in the cross-section, up to the first order, the covariance matrix behaves like a stock-time fixed effect and therefore will be absorbed by those in the saturated regressions.

2.2.3 Recovering Risk Aversion

Having discussed the identification of hedging demands and the way we estimate covariance matrices, we now turn to the issue of risk aversion. Let us first disregard any hedging demand for simplicity. The portfolio choice in this case takes the form of (2.4). Note that while we can separate changes in beliefs from changes in $\gamma_{i,t}$, the investor's risk aversion, we are unable to determine their levels. As a simple example, notice that $\tilde{\gamma}_{i,t} = 2 \times \gamma_{i,t}$ and $\tilde{\boldsymbol{\mu}}_{\boldsymbol{i,t}} - r_f\mathbf{1} = 2 \times (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f\mathbf{1})$ would yield the exact same portfolio choice as that implied by $\gamma_{i,t}$ and $\boldsymbol{\mu}_{i,t} - r_f\mathbf{1}$. In Section 2.6, we impose a plausible restriction on the *level* of subjective expected returns and risk aversion, namely, that fund managers expectations are formed in such a way to minimise the difference with ex-post realised returns[12]. Start from the following identities:

$$\begin{aligned}
\boldsymbol{r}_{t+1} - r_f\mathbf{1} &= \mathbb{E}_t[\boldsymbol{r}_{t+1} - r_f\mathbf{1}] + \boldsymbol{\epsilon}_{t+1} && (2.10)\\
&= (\boldsymbol{\mu}_{i,t} - r_f\mathbf{1}) + \boldsymbol{\epsilon}_{i,t+1} && (2.11)\\
&= \frac{\gamma_{i,t}}{\gamma_{i,t}}(\boldsymbol{\mu}_{i,t} - r_f\mathbf{1}) + \boldsymbol{\epsilon}_{i,t+1} && (2.12)\\
&= \gamma_{i,t}(\Sigma_t \boldsymbol{w}^*_{i,t}) + \boldsymbol{\epsilon}_{i,t+1} && (2.13)
\end{aligned}$$

The first line of the above expression is a definition for $\boldsymbol{\epsilon}_{t+1}$: realised returns have to be equal to expected returns plus an orthogonal prediction error. In the second line, we assume that the subjective expectation $(\boldsymbol{\mu}_{i,t} - r_f\mathbf{1})$ and the error $\boldsymbol{\epsilon}_{i,t+1}$ made by the investor are orthogonal. This can be interpreted as a requirement that the expected return is consistent with the law of iterated expectations [13]. The third line multiplies and divides this expectation by the investor's risk aversion $\gamma_{i,t}$. In the empirical counterpart, this will require that the instantaneous relative risk aversion is known to the manager at time t. Finally, we use equation (2.5) to rewrite (2.12) as (2.13). We can, therefore, pin down the level of risk aversion $\gamma_{i,t}$ by running multiple regressions across managers and/or time of stock realised returns on scaled portfolio weights. For instance, if we think that risk aversion is a

[12]Conditional expectations are the best predictor in a mean square sense, i.e., given the information set $\mathcal{F}_t$ and the random variable y_{t+1}, the conditional expectation $\mathbb{E}[y_{t+1}|\mathcal{F}_t]$ minimises $\mathbb{E}[(y_{t+1} - f_t)^2]$ over all the $\mathcal{F}_t$-measurable functions f_t.

[13]To see this remember that, according to our notation, the expected excess return of manager i using his information set at time t is $\mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}] = \boldsymbol{\mu}_{i,t} - r_f\mathbf{1}$. We can therefore rewrite (2.11) as $\boldsymbol{r}_{t+1} - r_f\mathbf{1} = \mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}] + (\boldsymbol{r}_{t+1} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}])$. If the law of iterated expectations applies under manager i's expectation, i.e., if $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}]] = \mathbb{E}_i[\boldsymbol{r}_{t+1} - r_f\mathbf{1}]$, it is easy to show that:

- $\mathbb{E}_i[(\boldsymbol{r}_{t+1} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}])] = \mathbf{0}$, i.e., there is no unconditional bias,
- $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}](\boldsymbol{r}_{t+1} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r}_{t+1} - r_f\mathbf{1}])'] = 0_{N\times N}$, i.e., the perceived expected return and the error are uncorrelated.

manager-specific quantity we can run the following regression:

$$r_{j,t+1} - r_f = \alpha_i + \beta_i (\Sigma_t \boldsymbol{w}^*_{i,t})_j + \varepsilon_{i,j,t+1} \tag{2.14}$$

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand for stock j for manager i, at time t scaled by the conditional covariance matrix Σ_t. The estimate for α_i will then be a measure of the bias or residual hedging demand. If $\alpha_i = 0$, i.e., the bias or hedging demand is not statistically different from zero, we would be able to interpret the estimate for β_i as the average coefficient of relative risk aversion of manager i, that is $\beta_i = \gamma_i$. It is important to notice that, while it might be interesting to pin down the *levels* of risk aversion and beliefs of each manager, the identification of the learning parameters comes from differential *changes* in beliefs in the cross-section of stocks held, hence it is not affected by our choice of the risk aversion parameter.

2.3 Data and Summary Statistics

In this section we describe the data that we use in the empirical analysis. Data on mutual funds and mutual fund managers' information are obtained from the Center for Research on Security Prices (CRSP) Mutual Fund database[14]. Given that we aim to conduct our analysis at the fund manager level, as opposed to the fund level, we need to construct a dataset of managers' careers. To do this, we first obtain a list of the managers that at any point in time are managing at least one equity fund. We then split each occurrence of multiple managers managing a fund at the same time into separate observations. We also disregard all the cases in which no manager name is available and all the observations where we have words such as "team", "group", "partners" or others that do not allow us to infer who was managing the fund. The most challenging part, however, is to account for cases in which a typo in the fund manager's name causes CRSP to treat the same manager as two different individuals. As an illustration, an individual named John Smith could, for exam-

[14] University of Chicago. Center for Research in Security Prices, I. (1960).

ple, appear as "John Smith", "J. Smith", "J Smith" or just "Smith". In order to tackle this issue, we first match names into pairs using a string matching algorithm. We match similar names using three different string distances: the cosine, Jaccard and Jaro-Wrinkler metrics, and we apply rather large distance-specific thresholds that allow us to keep the names which are sufficiently close. We subsequently proceed by manually checking the matched results which amount to more than 15,000 pairs of matched names. Out of these pairs, our manual exercise leaves us with roughly 20% of real matches which suggests that we are quite flexible with the distance thresholds. It is important to stress that, although our manual check might contain some errors, i.e., false positive matches and/or false match rejections, so long as these mistakes are random they only introduce noise in our estimates and cause no bias. More details on the process are provided in Appendix B.4. After matching the names we assign a unique index to each manager in order to build their careers. This exercise leaves us with 3,214 unique managers in our sample. We next match the above managerial data with CRSP mutual fund data based on the first and last date when a manager has been managing a given fund. We remove index funds, fixed-income funds and funds which mainly own foreign equities following Evans (2010), Benos et al. (2010) and Kacperczyk et al. (2006)[15]. We then match the fund information with mutual fund holdings data from the Thomson-Reuters Institutional Holdings database, using Russ Wermer's MFLinks tables. We finally merge the above data with CRSP data on stock returns and risk-free rates and Compustat-Capital IQ data on firm fundamentals[16]. Since we have monthly mutual fund and return data while holdings data are only available on a quarterly basis, we compute quarterly stock returns from the CRSP monthly data and proceed by merging with Compustat quarterly data. The final dataset comprises of over 13 millions observations for 3,214 distinct managers in the period 1980-2015[17]. Table B1 provides descriptive statistics. The first panel reports summary statistics regarding average and median past returns experienced by managers. As one should ex-

[15]Details on the removed funds can be found in Appendix B.4.

[16]Standard & Poor's Compustat Services, I. (1962).

[17]The number of observations includes a sizeble fraction of holdings that have zero weights but are included because they are part of the manager investment universe. The investment universe is constructed similarly to Koijen and Yogo (2019).

pect, past experienced returns tend to be right skewed with mean average returns that are larger than mean median returns (2.4% and 1.4%, respectively). While the standard deviation of average experienced returns is similar to the one of median experienced returns (10% and 11% respectively), counterintuitively, the latter seem to be more dispersed, implying that negative experienced returns tend to be right skewed (so that the median is smaller than the average) and positive experienced returns tend to be left skewed (so that the median is larger than the average).

The second panel of Table B1 regards expected returns, which are computed as explained in Section 2.2.1. In the rest of the paper we provide six measures of expected excess returns which we denote (1)-(6). The first issue regards the inclusion of zero weights[18]. Measures (1)-(3) include only positive weights, while measures (4)-(6) do include the zero weights[19]. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$. It is clear from the table that the measures are quite similar in terms of summary statistics. All the measures have an average expected excess return of about 1% per quarter and a median expected excess return of about 0.6%. It should also be noted that, while we have about 12.7 million data points if we consider the zero weights, the number of observations drops to about 5.4 million once we remove the zeros. Figure B1 sheds light on the sources of variation in beliefs. We provide a decomposition of the variation in expected excess returns according to measure (1) by regressing it against various fixed effects. Manager and stock fixed effects explain a small fraction of excess returns (11.63% and 14.20%, respectively), while time fixed effects explain more than half (55.73%) of the variation. This suggests that manager and stock immutable characteristics are relatively less important than aggregate time-varying factors in the formation of expectations. When we separately include manager, stock and time fixed effects the explanatory power rises to al-

[18]Similarly to the present paper, Koijen and Yogo (2019) discuss how the analysis might be affected by including or excluding zero weights.

[19]It might be important to know whether zero weights arise by choice or because the manager cannot short sell stocks that would otherwise appear with negative weights. Appendix B.3 shows how the optimal choice of a manager is affected by short selling constraints and how to deal with them when trying to recover beliefs.

most seventy percent (68.21%). If we allow for interactions between fixed effects, we can explain up to almost ninety percent (89.43%) of the variation in expected excess returns when we include manager-time and stock-time fixed effects. From this decomposition we learn that the largest part of the changes in expected returns is due to time-varying factors, then stock-specific characteristics and, finally, factors related to the manager. The addition of manager-time and stock-time fixed effects will remove the greatest majority of the variation in expected excess returns and will, thus, ensure that the results are driven by idiosyncratic variation in expected returns unexplained by systematic factors. This gives more credibility to our identification strategy.

Finally, we consider the data related to the managers' careers which can be analysed with the help of the last panel of Table B1 and Figures B2 and B3. The upper panel of Figure B2 provides information regarding the experience of the managers in the sample. We plot the number of managers by the first time they appear in the sample, which we call the starting date of the fund manager and denote it by $t_{i,0}$. The sample extends from 1980 to 2015 and covers a period of 35 years. Notice, however, that there are fewer managers who start their career in the first ten years compared to the rest of the sample. This can be attributed to low data coverage during the 1980s. Most of the managers in our sample begin their career in the late 1990s. We can observe, however, a wide range of manager starting dates up until the last sample year. We then proceed to construct a tenure variable which measures how many quarters have passed since the start of the manager's career, i.e., for a given manager i and date t, $\text{tenure}_{i,t} = t - t_{i,0}$[20]. The lower panel of Figure B2 displays the number of managers with a given level of accumulated tenure over the sample period, i.e., the empirical distribution of $(t - t_{i,0})$ for all i, t. Most of the managers in our sample are relatively young and inexperienced, but again, there is quite a large variation in tenure as well, ranging from less than a year up to some managers that are present in the whole sample (i.e., for a period of 35 years). Note that, by construction, the number of observations with a given level of accumu-

[20]Notice that for each manager we disregard the first quarter of experience, i.e., $t_{i,0}$, when computing the statistic.

lated tenure should be decreasing as, for example, a manager who has 5 quarters of accumulated tenure must also have accumulated 4 quarters of experience previously. In practice, this could be violated for two reasons: the first reason is that mutual funds were required to report holdings at a semi-annual level up until 2003 and only later regulators enforced quarterly reporting, as a result, some funds used to report holdings on a quarterly basis while others did so only on a semi-annual basis prior to 2003; second, there might be some missing data in our sample which means that we might be able to observe a given manager's career and holdings in a particular quarter but not in the previous one. The bottom panel of Table B1 shows that the average tenure is of 26.9 quarters (almost 7 years), but because of the positive skewness manifested in Figure B2, the median tenure is of only 22 quarters (5.5 years). We then proceed to the main object of interest of the paper, which is the relationship between each manager and stock. Figure B3 describes the relationship between fund managers and individual stock holdings. The first panel displays the date when a given stock-manager pair has first appeared in our sample which we call the starting date. For each manager i and stock j we can denote the starting date as $t_{i,j,0}$. Unsurprisingly, the largest number of such initiations have occurred in the late nineties and early 2000s, i.e., when the number of managers in our sample significantly increases. There is, however, large variation in the stock-manager starting dates which we exploit as part of our identification strategy. To see this, the second histogram depicts the length of the personal experience of a given manager with a given stock, i.e., for each manager i, stock j and date t, $\text{experience}_{i,j,t} = t - t_{i,j,0}$. It is clear from the histogram that there is a large variation in experience. The third panel of Table B1 shows that it ranges from 1 to 139 quarters, with a standard deviation of about 12.9 quarters. The standard deviation is of similar magnitude compared to the average (about 13.2 quarters) and the median experience (9 quarters). The main hypothesis of the paper is that this variation in stock-specific experience is associated with a variation in expected returns across managers. Finally, we can look at the maximal experience achieved for each stock-manager pair, in the bottom panel of Figure B3 and Table B1[21]. While the

[21]For each manager i and stock j, the maximal experience is defined as $\text{max. experience}_{i,j} =$

average maximal experience and its standard deviation are similar to the above (13.9 and 12 quarters respectively), the median maximal experience is larger (11 quarters, compared to 9 quarters of experience).

In the next section, we present the reduced-form results of our empirical analysis.

2.4 Reduced-form Results

The main hypothesis of the paper is that past experienced returns affect expected future returns. Moreover, if that is the case, we would like to further explore whether certain periods carry more relevance than others. In what follows, we show that differential stock-specific experience across managers indeed matters in the formation of expectations and, in particular, differences in the first and the most recent few quarters of experience play the most crucial role.

The empirical specification in this section relies on the following argument: we conjecture that the manager will try to estimate future returns by looking at the returns he has experienced over his career. A manager i with $T_{i,j,t}$ quarters of experience with a given stock j at time t might use the average experienced return as a sufficient statistic when forming expectations, i.e., his expected return for that stock can be represented as:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \bar{r}_{i,j,t} = \beta \left(\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} r_{j,t+1-k} \right) \tag{2.15}$$

where $\bar{r}_{i,j,t}$ denotes the equal-weighted average of stock j returns observed over the investor's experience horizon. Notice that the variation in the length of past experience $T_{i,j,t}$ allows us to exploit the cross-section of managers holding a given stock j as our source of differential treatment[22]. The coefficient β captures the average effect that past observed returns have on expectations formation, while the implicit constant weight $\omega_k = \omega = \frac{1}{T_{i,j,t}}$ means that all past observations are equally-

$\max_t\{\text{experience}_{i,j,t}\}$.

[22]On the other hand, the variation in the length of past experience $T_{i,j,t}$ for a given manager i at time t across different stocks is what helps us in disentangling preferences from expected returns.

weighted. This choice implies that investors attach equal importance to all observations, however, as the length of experience grows every observation receives a progressively lower weight. Note that this approach does not restrict managers from incorporating other sources of information in their estimation. This can be easily taken into account by saturating the regression with the proper controls. To reiterate, fixed effects, for instance, would account for the information or characteristics that all managers, or all stocks in the portfolio of a given manager, have in common; the coefficient on the average experienced return would thus provide a measure of the incremental effect of experience[23]. We, therefore, show in Table B2 the results of the following regression:

$$\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.16}$$

where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the previously defined equal-weighted average experienced return[24], $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To better disentangle the effect of experience we focus on the subsample of single-managed funds[25]. The results in the table confirm our main hypothesis: having experienced an increase of one standard deviation in average quarterly return leads to an increase in the expected excess return of between 0.103% and 0.151%; the results are both economically and statistically large and display very minor variation across specifications. This validates our intuition that the estimation method for the covariance matrix is not very consequential. Similarly, the inclusion of the zero weights has no effect on our main findings, even though the drop in R-squared shows that the zeros are indeed informative and cannot be fully explained by the fixed effects alone. The within R-squared shows that the average experienced returns explain between 0.6% and 0.9% of the variation in expected

[23]Notice that this implies that managers could very well use all past realised returns when they form expectations and this would be absorbed by the stock-time fixed effects. In particular, β would then measure the relative over-weighting of experienced returns.

[24]All the regressions in the paper use standardised explanatory variables for ease of interpretation.

[25]Section 2.4.1 analyses the case of co-managed funds, showing indeed that most of the effect washes out when we aggregate across managers.

returns. While this might seem low, it is in fact in line with the findings of Koijen and Yogo (2019) that observable characteristics explain a small part of the variation in investors' demands which is mostly explained by latent factors. Table B11 in Appendix B.6 reports the results of a similar regression with manager-time and stock fixed effects, and a number of time-varying stock characteristics, namely, profitability, investment, book-to-market ratio, market equity, and dividend-price ratio. The findings are similar in magnitude and statistically significant, and show that the effect of experienced returns is almost an order of magnitude larger than that of other known characteristics, confirming again the findings of Koijen and Yogo (2019) that standard predictors have a hard time explaining portfolio choices[26].

So far, we have assumed that the effect of experience is homogeneous. Alternatively, we could allow for more flexible weights in order to investigate whether certain periods matter more than others. Consider the following modified weight: $\omega_k = \frac{\delta_k}{T_{i,j,t}}$, such that $\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \delta_k = 1$. Namely, the manager estimates future returns from the *weighted* average of past experienced returns:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \sum_{k=1}^{T_{i,j,t}} \frac{\delta_k}{T_{i,j,t}} r_{j,t+1-k} = \sum_{k=1}^{T_{i,j,t}} \beta \delta_k \frac{r_{j,t+1-k}}{T_{i,j,t}} = \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k \tilde{r}_{j,t+1-k} \tag{2.17}$$

The weighting term δ_k is a number centred around one measuring the relative over- or under-weighting of a given past observation. If $\delta_k < 1$, then returns observed k-periods ago are under-weighted, while if $\delta_k > 1$ they are over-weighted relative to the previous benchmark. The last equality in equation (2.17) shows that if we rewrite $\tilde{\beta}_k = \beta \delta_k$ and $\tilde{r}_{j,t+1-k} = \frac{r_{j,t+1-k}}{T_{i,j,t}}$, then we can run a regression on experience-adjusted returns and obtain:

$$\beta = \frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k, \quad \delta_k = \frac{\tilde{\beta}_k}{\beta} \tag{2.18}$$

that is, the average effect of past experience can be obtained as the average of the k coefficients $\tilde{\beta}_k$, while the relative weight assigned to the k-periods ago return is given as the ratio of the coefficient on the k-th term and the equal-weighted average

[26] We do not report results for median experienced returns which are virtually identical.

of all coefficients.

In practice, this approach breaks down if we have to deal with varying experience lengths $T_{i,j,t}$, as the number of regressors would change together with $T_{i,j,t}$. For this reason, we group past returns into buckets as a means of fixing the number of regressors. In our first such specification we divide the stock-specific experience of the manager into five non-overlapping buckets of equal length, $|\Delta T^q_{i,j,t}|$, with $q = \{1, 2, 3, 4, 5\}$[27]. Table B3 reports the results of the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.19}$$

for $Q = 5$ and where $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, 3, 4, 5\}$, is the average return in the q-th bucket. Table B4 reports the results for ten non-overlapping buckets of equal length, i.e., the specification in equation (2.19) for $Q = 10$. In both cases we focus on the subsample of single-managed funds. To better visualise the results, the estimated coefficients of a regression with five buckets are reported in the upper panel of Figure B4, while the bottom panel reports the results for ten buckets. The picture immediately reveals that the effect of past experienced returns is clearly neither constant nor monotone. Consider, for instance, our first model of expected returns with $Q = 5$ for which we show results in column (1) of Table B3: a one standard deviation increase in experienced average quarterly return in the most recent or in the earliest period of holding the stock increases the expected return by roughly 0.25% ($\beta_1 = 0.276$ and $\beta_5 = 0.238$); on the other hand, the effect of an increase of one standard deviation midway through the manager's experience has an effect lower by almost an order of magnitude ($\beta_3 = 0.041$). Figure B4 confirms that the effect of experienced returns is "U-shaped" regardless of whether we include the zero weights and independently from the estimator for the covariance matrix used. The lower panel of the figure reports the results for $Q = 10$, painting almost an iden-

[27]To cast this specification in terms of the previously discussed model, let us denote each bucket by $\Delta T^q_{i,j,t}$ and its length by $|\Delta T^q_{i,j,t}|$. We then have that $\delta_k = \beta_q \frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|}$, where for each time index k in bucket $\Delta T^q_{i,j,t}$ we assign a common effect β_q and take the average return $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} = \sum_{k \in \Delta T^q_{i,j,t}} \frac{r_{j,t+1-k}}{|\Delta T^q_{i,j,t}|}$. Notice that $\frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|} \approx 5$, where the approximation derives from the fact that we have to split ties when the experience length is not a multiple of five.

tical picture. The coefficients for ten buckets are similar in magnitude to those for the regression with five buckets and follow the same "U-shaped" pattern. We report in Appendix B.6 the results for various other specifications: Tables B12 and B13 report the results of the previous models with stock fixed effects and the previously mentioned controls, while Tables B14 and B15 describe the results for a model with three non-overlapping equal-sized buckets; finally Tables B16, B17 and B18, B19 report the results for three non-overlapping buckets of unequal length (with stock-time fixed effects or stock fixed effects and varying controls), where the first and last buckets consist of four and eight quarters, respectively. All these specifications confirm the previously discussed results: experienced returns are important in determining expected returns and most of the impact comes from most recent and earliest stock-specific observations. This is evidence in favour of the so-called serial-position effect, concept well studied among researchers in psychology (Murdock, 1974). Moreover, our findings reconcile two apparently distinct phenomena observed in previous research: on the one hand, Malmendier and Nagel (2011) show that economic agents are principally affected by recent experience, while on the other hand Kaustia and Knüpfer (2008), Hirshleifer et al. (2021) and Hoffmann et al. (2017) report evidence in favour of the *primacy effect* or *first impression bias*. We show that both effects are present in mutual fund managers and that they need to be separately considered.

2.4.1 Co-managed Funds

So far we have focused our attention on single-managed funds, but one might be interested to know whether the above findings are, in fact, weaker when managers work in teams. In this section we check the impact of the number of managers within a team on the effect of experience. Our hypothesis is that personal stock-specific experiences should partly offset each other within a team, so long as the managers that form part of the team have followed different career paths. To explore this hypothesis, we run the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.20}$$

We split managers into subsamples based on the number of co-managers they work with, i.e., $n_{i,t} \in \{1, 2, 3, 4 \text{ or more}\}$ signifies that the manager works in a team of one, two, three or four or more people. We thus obtain a different set of coefficients $\beta_{q,n}$ for each combination of buckets and size of the management team. We report in Table B5 the results of this exercise for $Q = 5$ buckets; the results for the specification with 10 buckets are reported in Table B20 in Appendix B.6. To better visualise the results, Figure B5 displays the coefficients $\beta_{q,n}$. The two plots on the left-hand-side of the figure show the results for measure (1) while the right-hand-side plots display the coefficients for measure (4). The first row reports the results for $Q = 5$ and the bottom row for $Q = 10$ buckets. As one can see in Table B5, the coefficient on the most recently experienced returns for single-managed funds is more than twice as large as the same for funds managed by two managers; the difference in coefficients is even larger for the earliest bucket of returns, more specifically, the effect of returns observed at the beginning of a stock-specific experience is more than ten times greater for single-managed funds compared to funds managed by at least two people. The effect on managers working in teams of three or more is orders of magnitude lower, while still statistically significant for recent experienced returns. On the other hand, the effect of early returns loses significance. The above is visually confirmed by the plots in Figure B5 showing a rather steep decrease in the coefficients on the earliest bucket of returns across teams of different sizes, especially when going from a single-managed fund to a fund managed by two professionals. The findings are equally pronounced for the specification with ten buckets.

The results seem to suggest that a considerable part of personal experience washes out in the cross-section of managers working in the same team, and more so the further we go in the past since managers are more likely to change teams over a longer period of time. On the other hand, recent returns affect all co-managers in a similar way as they have presumably gone through the same recent experience, having been working for the same fund. This could justify the difference in spreads observed between buckets at different horizons, especially if we compare single-managed funds with those managed by two individuals.

2.4.2 Taxes

In what follows, we investigate the impact of taxes on managers' investment decisions and the potential explanation the tax regime might have thereof. More specifically, we examine whether tax considerations can absorb the effect that past experience has on portfolio weights and expectations formation. The differential treatment of short-term and long-term capital gains in terms of their taxation, together with the possibility to offset capital gains with capital losses, suggests that mutual funds will try to defer the realisation of gains and accelerate the realisation of losses. This implies that it is optimal from the point of view of minimising the tax bill for mutual funds to hold on to assets that performed well in the past and sell assets that had subpar performances[28].

This, in turn, means that the previous results could be simply driven by tax considerations. One way to solve the problem is to model the optimal selling decision in the spirit of Barclay et al. (1998) or Sialm and Zhang (2020) and check if the effect of experienced returns survives after we have taken tax considerations into account. However, in what follows we take a reduced-form approach and make use of the large amount of data on managers who have managed different funds in their career. In particular, we focus on the subsample of manager-stock pairs where the manager had positive holdings of the stock in the past while managing a different mutual fund compared to the one that he is currently working for. In this setting, tax considerations should be muted given that capital gain overhangs cannot be transferred from one fund to another.

Table B6 reports the results of a regression of expected returns on five buckets of

[28] Bergstresser and Poterba (2002) show that inflows to mutual funds, and therefore managers' compensation, are affected by the amount of unrealised capital gains, implying that there might be a tension between postponing capital gains indefinitely to provide better after-tax returns for current investors and attracting new investors. Barclay et al. (1998) explicitly tackle this question, showing that indeed managers tend to realise gains early to attract new investors.

past experience for only those managers that have changed funds, while Table B7 reports the results when we split the previous experience in ten buckets. The results are then summarised in Figure B6 where the upper panel reports the results for five buckets and the lower panel for ten buckets. While the number of observations is greatly reduced (from about 800,000 to slightly more than 110,000 observations if we do not include zero weights, and from about 2 million to approximately 225,000 if we do), the economic and statistical significance of the coefficients is virtually unchanged confirming the previous findings: experienced returns have a sizeable influence on expected excess returns, with the majority of the effect coming from the extreme buckets. If, for instance, we consider measure (1) we notice that the coefficient on the most recent bucket goes from 0.276 to 0.224, while on the earliest one from 0.238 to 0.199. We infer, therefore, that no more than 20% of the effect might be due to tax considerations and we confirm both the *recency* and the *first impression bias*.

Having presented the reduced-form results of our analysis, we now develop a simple three-parameters model of learning and proceed with its estimation.

2.5 Parametric Estimation

The reduced-form evidence of the previous section teaches us that: experience matters, i.e., average experienced returns are an important determinant of expected returns and; the effect of experience is neither constant nor monotone, in particular, earliest and most recent experience matter the most. However, as shown in Section 2.4, estimating the shape of the weighting function requires us to drop a sizeable amount of observations and potentially lose significant information. For this reason we now posit a functional form for the learning weights and try to estimate its parameters. As Figures B4, B5 and B6 show, we need to allow for non-monotone weights if we want to accurately fit the data. Similarly to Section 2.4, we assume that the manager uses a weighted average of experienced returns in order to predict future returns. Recall the model in equation (2.17) where the weights $\frac{\delta_{i,j,t,k}}{T_{i,j,t}}$ capture the differential effect of returns experienced at different points in time. In this section we directly model these weights as follows:

$$\omega_{i,j,t,k} = \frac{\delta_{i,j,t,k}}{T_{i,j,t}} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}} \tag{2.21}$$

The functional form in equation (2.21) is similar to the one used by Malmendier and Nagel (2011) and Malmendier and Nagel (2016)[29]. The weighting function used in these papers depends only on $T_{i,j,t} - k$ and, as such, it confounds two separate effects: the *first impression bias* and the *recency bias*. On the other hand, our weighting function has the advantage of disentangling between these effects: the term $T_{i,j,t} - k$ measures the distance between the return observed at time $t + 1 - k$ and the beginning of a stock-specific experience, hence capturing the *first impression bias*, while k measures the distance from the current date t, thus capturing the *recency bias*. Figure B7 shows how flexible the parsimonious parametrisation introduced in equation (2.21) is. We plot in blue the weighting function for a manager with $T_{ij,t} = 50$ quarters of experience for all the combinations of $\{\lambda_1, \lambda_2\} \in \{-0.1, 0, 0.1\} \times \{-0.1, 0, 0.1\}$[30] and compare it to the black dashed line representing the benchmark $\frac{1}{T_{i,j,t}}$ where the manager equally weights each observation that forms part of his experience. The first parameter, λ_1, governs the strength of the *first impression bias*: when it is negative, the manager is overweighting early experiences relative to the benchmark scenario. The second parameter, λ_2, controls the strength of the *recency bias*: when the sign of λ_2 is negative the manager overweights recent observations relative to the benchmark, and vice versa. As one can see from the examples in Figure B7, using only two parameters we are able to capture a variety of shapes including linear, convex or concave, increasing or decreasing, monotone or non-monotone weighting schemes arising from the interplay of the *recency* and *first impression bias*. Given the evidence from the reduced-form regressions we expect λ_1 and λ_2 to be negative, implying that the managers are subject to both effects. Similarly to the model in equation (2.19) we include manager-time and stock-time fixed effects to get rid of potentially time-varying unobservable characteristics shared across stocks and managers, respectively. Ta-

[29]Our weighting scheme collapses to the one used by Malmendier and Nagel (2011) when $\lambda_2 = 0$.

[30]Figure B12 in Appendix B.6 plots the weighting function for $\{\lambda_1, \lambda_2\} \in \{-2, 0, 2\} \times \{-2, 0, 2\}$.

ble B8 reports the NLS estimates of the following regression[31]:

$$\mu_{i,j,t} - r_f = \beta \left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k} \right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.22}$$

$$\omega_{i,j,t,k} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}$$

Consistent with the reduced-form evidence, both λ_1 and λ_2 are negative and statistically significant across all specifications. The magnitude of the effects is illustrated in Figure B8 where we plot the weighting functions at median and average experience of $T_{i,j,t} = 9, 13$ quarters, respectively, using the empirically estimated values for λ_1 and λ_2 under model (1). It is evident that the weighting function is always convex and non-monotone, implying that managers overweight the most recent and the earliest returns observed; for instance, a manager with an experience of nine quarters will assign a weight of 0.204 (0.347) to the most recent (earliest) observation, which is 1.84 (3.13) times the benchmark of 1/9. On the contrary, he will only assign a weight of 0.043 to the middle observation which is 0.39 times the benchmark weight. The results display a slight asymmetry with λ_1 being always larger in magnitude than λ_2 implying that the *recency bias* is marginally weaker compared to the *first impression bias*. This is, however, not a robust feature of the data: Table B21 in Appendix B.6 shows that λ_1 and λ_2 are almost identical once we include only manager-time and stock fixed effects, indicating that a large fraction of the recency bias might be captured by stock-time fixed effects as we should expect. Pinning down the actual magnitude of the two biases is extremely difficult given that we have to get rid of a large fraction of the variation in expected returns to achieve identification. Finally, the parameter β in Table B8 measures the average impact of past experience on expected excess returns: the estimates range between 0.139 and 0.207. This is about 4 basis points larger than the baseline results in Table B2 where we do not allow for varying weights[32]. We therefore confirm that once

[31] Appendix B.5 provides more details on the estimation procedure.

[32] Note that all the results presented refer to standardised variables. In the case of the results in this section we estimate β and then scale its value by the standard deviation of $\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k}\right)$. This is to avoid directly scaling the weighted average which would affect the computation of the

we take into account the possibility that recent and early returns might have a differential impact, we find an incremental effect of experience on expected returns.

2.6 Risk Aversion

As explained in Section 2.2.3, our methodology allows us to examine in more detail the preferences of investors. Recall equations (2.10)-(2.13); if we assume that subjective expected returns obey the law of iterated expectations, we are able to extract the risk aversion of managers by exploiting the cross-section of individual stock holdings. Running regressions of realised excess returns on scaled demands, as shown in equation (2.14), we can obtain an estimate for the risk aversion parameter γ and the bias (or residual hedging demand). We start this section by providing evidence from pooled regressions and then proceed to show results pertaining to the distribution of γ_i obtained from multiple regressions. Table B9 reports the results of the following pooled regression:

$$r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i,t}})_j + \epsilon_{i,j,t+1} \tag{2.23}$$

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i,t}})_j$ is the scaled demand for stock j of manager i, at time t. If we assume that preferences are constant across managers and time, we obtain a risk aversion coefficient close to unity (between 0.915 and 1.283 across specifications) for our *representative investor*. While the estimate is low compared to other measures obtained from equity returns (Mehra and Prescott, 1985; Kocherlakota, 1996), it is consistent with measures derived from labour choices (Chetty, 2006) and option prices (Martin, 2017). Our *representative investor* displays a quite large and statistically significant bias (or residual hedging demand) of about 1% per quarter.

The pooled results in Table B9 mask a sizeable amount of variability across managers. For this reason, we proceed by estimating separate regressions, one for each

gradient of the right hand side of equation (2.22) needed to obtain standard errors.

manager in the sample:

$$r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1} \tag{2.24}$$

Given that there seems to be limited difference resulting from the choice of the covariance matrix Σ_t, we report the results using the sample covariance $\hat{\Sigma}^1_t$. Table B10 reports summary statistics for elicited risk aversion and bias referring to measure (1)[33]. We obtain a median (average) relative risk aversion of 1.117 (1.236), in line with the pooled results; however, there is a wide dispersion in the estimates with a standard deviation of 5.850. The estimates display positive skewness and are leptokurtic. When we allow for variation in preferences across managers, the bias is reduced on average: the mean bias is only 0.7% and the median bias is 1% per quarter. Figure B9 displays histograms of the distribution of α_i and γ_i after we have removed outliers. Unfortunately our methodology does not prevent us from obtaining negative values for γ_i whenever the cross-section of revealed beliefs is negatively correlated with realised returns. Most of the mass, however, seems to fall in the positive value region.

We then proceed to exploit the variation of preferences across managers and analyse whether tenure affects risk aversion and bias. Figure B10 displays the bias and the risk aversion as a function of tenure for measures (1) and (4). It is hard to detect a specific pattern in either of the measures; longer tenures seem to be dominated by noise, given that they make use of fewer estimations by construction. Finally, Figure B11 reports the results by date: also in this case it is hard to detect any conclusive evidence. Unfortunately, our measures of risk aversion cannot be used to predict or explain future returns given that they have been obtained from them: by construction they represent the best linear predictor of $r_{j,t+1} - r_f$ given the information contained in $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$.

[33]The results for measure (4) can be found in Appendix B.6.

2.7 Conclusions

This paper contributes to the literature on the effect of personal experience on learning and expected returns by analysing a large sample of more than 3,000 professional investors (mutual fund managers) that have been tracked throughout their careers in the 35 years period between 1980 and 2015. Section 2.2.1 shows that in a variety of cases it is possible to invert the portfolio demands of our investors to obtain their subjective expected returns by using the identifying assumption that, while beliefs vary at the stock-investor-time level, risk aversion varies at the investor-time level, i.e., risk aversion is constant in the cross-section of stock holdings of a given manager. Similarly, we are able to account for many cases in which demands display a hedging component by saturating the regressions with fixed effects. Indeed, as we show in Section 2.3, almost ninety percent of recovered expected returns can be explained by manager-time and stock-time fixed effects. We then provide reduced-form evidence that professional investors overweight experienced returns compared to other information shared across stocks and individuals: having experienced a one standard deviation increase in quarterly returns on average leads to an increased expected return of about 10-15 basis points per quarter. Various reduced-form specifications in Section 2.4 and the parametric estimation in Section 2.5 confirm that the effect of experienced returns is neither constant nor monotone. We show that investors exhibit *recency* and *first impression bias*: an investor with a stock-specific experience of nine quarters overweights the most recently observed quarterly returns by 1.84 times and the earliest experienced returns by 3.13 times relative to the constant weight benchmark. These results are most apparent for managers working alone, as opposed to in a team of two or more, suggesting that a significant fraction, though not the entirety, of the effect of personal experience cancels out once aggregated. By looking at managers who have switched funds, we eliminate the possibility that these findings are purely driven by tax considerations: more than 80% of the effect remains unexplained by tax concerns. We finally turn to the issue of estimating risk aversion and find that a *representative investor* displays a coefficient of relative risk aversion around unity.

The paper also finds that individual investors exhibit biases when forming expectations. Finally, when we look at more disaggregated measures, we find that there is a large heterogeneity in biases and risk aversion across time and investors. The results in the paper can inform theorists willing to model the preferences and the learning behaviour of professional investors in a way that is consistent with the evidence obtained from portfolio holdings. Consistent with theory, more than half of the variation in expected excess returns can be explained by a common time varying component. However, an incremental forty percent is due to investor-specific and stock-specific time-varying effects, hinting at the possibility of time variation in preferences and stock-specific factors shared across investors. Finally, if interested in modelling the idiosyncratic part of expected returns, one should pay particular attention to behavioural factors which play a prominent role as shown by the evidence provided in this paper.

3. Living on the Edge: the Salience of Property Taxes in the UK Housing Market

FRANCESCO NICOLAI, MARCO PELOSI AND SIMONA RISTESKA[1]

A standard tenet of economic theory is that the statutory incidence of taxes is irrelevant for their economic incidence[2]. It should also be the case that whether a tax is paid at the moment of transaction or later is irrelevant for its incidence, as long as we take into account the time value of money and the riskiness of the cash flows. By looking at the UK residential property market, this paper shows that this is not the case and that deferred taxes have a markedly lower incidence compared to taxes paid at the time of decision-making.

Together with France, the United Kingdom is one of the few countries receiving a sizeable fraction of revenues from property taxes, amounting to about 4.3% of GDP or more than £84 billion in 2016 (European Commission (2018)). The two main taxes levied on domestic properties are the Stamp Duty Land Tax and the council tax. The former is a tax levied on the transaction value of land and any buildings and structures thereon. The fact that its statutory incidence falls on the buyer, who is required to pay the tax liability to the HM Revenue and Customs within very few weeks from the completion of the transaction, and the fact that the tax represents a lump sum ranging between 1% and 7% of the property value are features that make the stamp duty tax particularly salient at the moment of purchase. The latter, which is the focus of the present paper, is a tax levied by the local government on a yearly basis. The council tax is levied on the resident, as opposed

[1]We are grateful to Vicente Cuñat, Daniel Ferreira, Dirk Jenter, Christian Julliard, Daniel Paravisini, Andrea Tamoni, Michela Verardo and the participants at the LSE PhD seminar for the useful comments on the paper. We thank Vittorio Raoul Tavolaro for invaluable research assistance. The paper contains HM Land Registry data © Crown copyright and database right 2019. The data is licensed under the Open Government Licence v3.0. 1. We thank the University of Glasgow - Urban Big Data Centre for providing Zoopla property data. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[2]Kotlikoff and Summers (1987) provide a detailed review of classical theory on tax incidence.

to the house owner, and is based on the property value in 1991. While the council tax is extremely salient at the moment when it needs to be paid, we show that this is not the case at the moment when properties are purchased even though, in present value terms, it is similar to or even larger than the stamp duty tax. By using the geographical discontinuity at the border of different local authorities in the London area, we are able to estimate the incidence of the council tax on property prices and contrast it with the incidence of the stamp duty tax estimated, among others, by Best and Kleven (2018). The London area is particularly suitable for the estimation because of the sharp nature of the council borders and the large dispersion in council tax rates across Boroughs. For instance, Figure C1 depicts a road that is at the border of the Borough of Westminster and the Borough of Kensington and Chelsea. As can be seen from the picture, the houses on both sides of the street are otherwise identical except for the fact that they pay quite different council tax amounts: the ones on the left pay £2, 279 per year in council tax while those on the right pay £1, 421 per year. If we discount the future payments as a perpetuity at a rate of 4%, similar to the mortgage rates observed in sample, we obtain that the difference between the two present values amounts to £21, 450 (about \$28, 000). The tax differentials become even more significant once we consider the fact that many London Boroughs share services, such as waste management, and that many other amenities, such as access to parks, schooling and religious facilities, are not strictly limited to residents of a given Borough. In Section 3.3 we show that the price of similar properties on opposite sides of a border does not adjust for differentials in council tax amounts. By employing a variety of estimators, we establish that the council tax incidence is never statistically negative. We then proceed in Section 3.4 to set up a model where downpayment-constrained households purchase a house and pay two sets of taxes: a lump sum stamp duty tax levied at the moment of purchase and a periodic council tax. We move on to perform a Bayesian analysis in Section 3.4.1 where we provide a posterior range for council tax incidence using priors that are economically motivated. In all these estimates, the incidence of council tax on property prices is too low relative to existing estimates of the incidence of other property taxes, even after accounting for time value of money and the fact that discount rates might be larger because of borrowing constraints. These findings can be rationalised in a model where agents neglect taxes that are levied in the future. We show in Section 3.4.2 that then a trade-off between the two types of taxes arises: the stamp duty tax is distortionary because agents are liquidity constrained; on the other hand, the council tax leads agents to over-consume the housing good and, therefore, distorts their consumption choices by reducing available income. As a result, we demonstrate that the Government can optimally tune the two taxes to minimise distortions for a given level of revenue.

The present paper adds on to the burgeoning literature on behavioural public finance and the salience of taxes (or the lack thereof). Chetty et al. (2009) is the first paper to empirically estimate how differences in salience can alter the behaviour of economic agents. They intervene in a grocery store in order to modify the salience of sales taxes and show that the incidence on buyers is largely reduced when taxes are made fully salient. In a second experiment they compare the effect of excises taxes, which are included in posted prices, and sales taxes, which are not explicitly included, on alcohol demand and again show that tax salience plays an important role in consumer behaviour. The setting in the present paper is quite similar to the second experiment in Chetty et al. (2009), given that the stamp duty tax is paid upfront while the council tax is deferred and thus less salient. For policy reasons, however, the question of property taxes is of greater importance because of the large amounts of money involved and the fact that it is very difficult for agents to learn since buying a new property is typically a once-in-a-lifetime event. Following Chetty et al. (2009), other papers have also explored the question of tax salience, for instance, Feldman and Ruffle (2015) and Feldman et al. (2018) have replicated the findings in laboratory experiments, while Finkelstein (2009) similarly shows that the introduction of electronic toll payments raises toll expenditures. Taubinsky and Rees-Jones (2018) further explore the topic by showing that there is large variation in the way agents react to tax salience and investigate policy implications. The present paper is also akin to Allcott (2011) who demonstrates that a similar bias is present in the automobile market, namely, car buyers fail to correctly price in the future energy cost at the time of purchase. As in Allcott (2011), our conclusions also rely on the choice of an appropriate discount factor. We show in Section 3.4.1 that the bias persists even after allowing for large discount rates. In a similar vein, using Norwegian data, Agarwal and Karapetyan (2016) explore the effect of non-salient debt features on households' purchasing decisions and show that they do not fully factor in the added cost. The authors show that the mispricing was eliminated once these features became fully salient. Finally, the paper extends the literature on property taxes; among others, we use the results of Besley et al. (2014) and, in particular, Best and Kleven (2018) to compare our estimates of the council tax incidence with their stamp duty incidence estimates in order to highlight the lack of salience of the former.

The rest of the paper is organised as follows: Section 3.1 describes the data and the institutional setting; Section 3.2 gives evidence of the geographical distribution of council taxes and points out that this can significantly bias our estimates if not appropriately taken care of, before proceeding with the details of our identification strategies; Section 3.3 presents the empirical estimates of the council tax incidence; Section 3.4 develops a stylised model to help interpret the findings and shows that the estimated incidence is too low to be consistent with fully-salient taxes, before exploring some policy implications; and finally, Section 3.5 summarises and concludes the paper.

3.1 Data

To estimate the incidence of council taxes we need access to data on property characteristics and house prices, as well as council taxes paid. Price paid data on house transactions are readily available from the HM Land Registry website. This dataset contains information about all residential properties transacted in England and Wales from 1995 that have been sold for full market value[3]. The dataset comprises of the price paid, the transaction date and, most importantly, the address of the house which allows us to pinpoint the exact location of every property. Ad-

[3]Data excluded from the dataset include commercial transactions, property transactions that have not been lodged in with HM Land Registry and transactions made below market value. For more details on the property sales not included in the dataset the reader can visit the HM Land Registry website: https://www.gov.uk/guidance/about-the-price-paid-data.

ditionally, the data provide us with information on the property type, which can be one of five possible categories (a detached, semi-detached, or terraced house, a flat/maisonette and other), the age of the property (classified into old or new to distinguish between newly built properties and already established buildings) and the duration of tenure, i.e., whether the property is under a freehold or leasehold[4].

Since we would ideally like to compare properties that are as similar to each other as possible, we need more information on property characteristics. For this purpose we make use of two additional datasets: the Zoopla Property data and Domestic Energy Performance Certificates. The Zoopla Property data[5] has been collected by Zoopla, one of the UK's leading providers of property data for consumers and property professionals, giving free access to information on 27,000,000 property records, up to 1,000,000 property listings and 15 years of sold prices data. The dataset covers the period between 1st January 2010 and 31st March 2019 for properties located in Great Britain (England, Wales, Scotland). The dataset contains details on characteristics such as property location, property type[6], whether the property has been categorised as residential or commercial[7], number of bedrooms, number of floors, number of bathrooms, number of receptions and whether the property is listed for sale or for rent[8]. In addition, we also have access to the asking price for both rents and sales, however, we use the more accurate transaction price from the HM Land Registry dataset. The second source of house characteristics comes from the Ministry of Housing, Communities and Local Government. On their website, one can access the Energy Performance Certificates (EPC) for domestic and non-domestic buildings. For domestic properties, before 2008 cer-

[4]Note that leases of seven years or less are not recorded in the dataset.

[5]The access to the dataset has been kindly provided by the University of Glasgow - Urban Big Data Centre. Access to the dataset for research purposes can be obtained directly through the Urban Big Data Centre. The data has been collected by Zoopla. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[6]Property types include: barn conversion, block of flats, bungalow, business park, chalet, château, cottage, country house, detached bungalow, detached house, end terrace house, equestrian property, farm, farm house, finca, flat, hotel/guest house, houseboat, industrial, land, leisure/hospitality, light industrial, link-detached house, lodge, longère, maisonette, mews house, mobile/park home, office, parking/garage, pub/bar, restaurant/cafe, retail premises, riad, semi-detached bungalow, semi-detached house, studio, terraced bungalow, terraced house, town house, unknown, villa and warehouse.

[7]We keep only properties categorised as residential.

[8]For the time being we only keep properties listed for sale.

tificates could be lodged on a voluntary basis. From 2008 onwards, however, it has become mandatory for accredited energy assessors to lodge the energy certificates. Consequently, the data coverage drastically improves around that time, as does our ability to match these data with the price paid data. More specifically, the matching rate jumps from about 50 percent to over 90 percent around 2008. The dataset contains information on the location, property type, total floor area, number of storeys, number of rooms, floor level and height, along with many indicators of energy efficiency and quality of glazed surfaces. The final piece of data needed to conduct our analysis is related to council tax data; in the following section we describe in more detail the functioning of this property tax and the relevant data.

3.1.1 Council Tax

The taxation of properties in the United Kingdom is peculiar compared to other OECD countries, representing a rather large source of both central Government and local authorities' revenues. The three main taxes levied on properties are the council tax, business rates and stamp duty taxes. Council taxes are levied on each occupier, rather than on the owner, of domestic properties. The tax is one of the few levies in Great Britain being both set and collected by local authorities (Boroughs in the case of London) and it represents one of their major sources of revenue (around one-third of total revenue), the other sources being commercial property taxes (business rates) and transfers from the central Government. The tax is based on a classification in eight bands (A-H) based on the value of the property as established by the Valuation Office in 1991; newly built properties are assigned to a band, after having their current value converted into the value of an equivalent property in 1991. Each London Borough is responsible for setting the annual tax amount to be paid by a property in band D every year; the amount to be paid by other bands is automatically set as a ratio to the amount for band D[9]. Bands C and D represent the largest fraction of dwellings (about 50 percent of the total), but there is variation across Boroughs with central properties being skewed towards

[9]The ratios are constant across Boroughs and are as follows: band A 6/9, band B 7/9, band C 8/9, band D 1, band E 10/9, band F 13/9, band G 15/9, band H 2.

higher valued bands compared to properties in outer Boroughs. Figure C2 shows the time series of the council tax payable per band per Borough. Each panel in the figure depicts the amount payable by different bands showing that, by construction, the tax moves in locksteps across bands. More interestingly, it should be noted that there is a wide dispersion in amounts payable across Boroughs, even though the ranking across different local authorities remains almost constant with the only exception being the Borough of Hammersmith and Fulham where taxes have been slashed starting from the late 2000s. After a marked increase in council tax rates in the early 2000s, the freeze mandated by the central Goverment after the 2008 financial crisis is visible in the time series; since 2011, taxes can be raised only by a centrally set amount unless a local referendum allows the authority to do so. We show in Section 3.2.1 that the geographical distribution of council tax rates is not random and could severely bias any estimate of incidence, given that central (and pricier) Boroughs tend to set lower council tax rates. This is mainly because central Boroughs tend to have larger fraction of properties in higher bands; for instance, the Borough of Kensington and Chelsea raises more than fifty percent of its revenues from bands G and H, while Barking and Dagenham raise less than five percent from these bands.

We obtain information on council tax band assignment from the website of the Valuation Office Agency, which provides data on the full address and the council tax band for each property in Great Britain. The average amount to be paid in each London Borough by each band in the period 1999-2018 is obtained from the London Datastore managed by the Greater London Authority.

In the following section, we provide some descriptive statistics of the data we have mentioned so far.

3.1.2 Descriptive Statistics

Figure C3 shows the distribution of transaction prices for domestic properties in London, truncated to exclude extremely high property prices which are, however, included in the analysis. The data consists of 889,925 observations in the period between 1999 and 2018 for which property characteristics and council tax information is available. We confirm that the distribution is highly skewed with the average and median property values being £366,528 and £250,000, respectively. It is immediately obvious that there is a large degree of bunching in prices, as noted for instance in Best and Kleven (2018). The bunching mainly happens just before stamp duty notches, which allows Best and Kleven (2018) to estimate the local incidence of this tax. Figure C4, for instance, shows the large extent of bunching at the threshold of £250,000 (upper panel) and £500,000 (lower panel) where the stamp duty tax jumps from 1% to 3% and from 3% to 4%, respectively. Best and Kleven (2018) estimate a rather large incidence of stamp duty tax on property prices and argue in favour of evidence of rather strict borrowing constraints; we use their estimates to inform our analysis of the incidence of the council tax, allowing us to disentangle how much of the incidence is due to borrowing constraints (or the lack thereof) and how much is attributable to pure time discount. Figure C5 shows the distribution of house prices per band. The vertical red lines depict the median price within each band. As one should expect, higher bands tend to have houses with higher average prices although there is a large dispersion within bands. This is because prices have increased a lot over the past twenty years, especially for more central and higher-banded properties. This makes it essential that we compare only transactions occurring in close periods. Moreover, one should notice that the number of properties belonging to bands C and D dominates the rest, as previously mentioned. In Figures C6, C7, C8, C9 and C10 we show that there is a wide dispersion of transaction prices based on house characteristics such as property type, number of rooms, property age and duration. There is a disproportionate amount of flats in our sample, which we see as an advantage in our estimation, as flats are much more likely to be similar to each other relative to other property types. Detached houses are most expensive, with a median price of £525,000, followed by semi-detached houses (£319,950) and terraced houses (£270,000), and finally, flats are the cheapest category (£195,000). Naturally, the house price is increasing in the number of rooms with the median value of each additional room being about £40,000 in the full sample. Newly-built properties represent a minority in our sample and trade at a small discount relative to established buildings. This is due to the geographic distribution of the housing stock in London where older properties tend to be in the more sought-after central areas. However, there is some heterogeneity when we look at the year of construction: properties built before 1949 sold at a median of £287,000 close to those built after 2003 (£275,000), while properties built in the period 1950-1982 and 1983-2002 sold at lower prices (£215,000 and £200,000, respectively). This pattern can be explained both by differences in type and location across groups. Finally, it can be noted that properties under a freehold ownership have a higher median price (£305,000) compared to leasehold properties (£195,000).

After having described the data, we proceed to the discussion of our empirical strategy in the next section.

3.2 Empirical Strategy

3.2.1 Evidence of Selection

The main issue that arises when estimating the incidence of council taxes is the fact that the cross-sectional distribution of council tax amounts across Boroughs is very strongly correlated with other characteristics that affect house prices. To see this, Figure C11 shows a map of the distribution of Band D council tax amounts payable for each London Borough along with the respective distribution of house prices. Panel C11a shows the distribution of council taxes in 2000, where taxes increase moving from yellow to red; Panel C11b the distribution of house prices in the same year, where prices increase moving from light blue to brown. Panel C11c shows the distribution of council taxes in 2018, while panel C11d the distribution of house prices in the same year. It is visually striking that councils with lower taxes tend to have higher house prices. For instance, the City of Westminster had the lowest Band D council tax in 2000 (£375.17) and the second highest average house price (£357,925), after the Borough of Kensington and Chelsea (£726,908) which had the fourth lowest council tax (£623.38). In 2018 the same holds true, with the

City of Westminster having the lowest council tax (£710.50) and the second highest average price (£1,612,231), after Kensington and Chelsea (£3,040,547) which had the fifth lowest council tax (£1,139.41). In general, it is clear from the map that Boroughs that lie further from the centre tend to have higher council taxes and lower prices, while the more central Boroughs tend to exhibit the opposite pattern. To confirm the intuition obtained from Figure C11, we can run a naïve regression of house prices on house characteristics and council tax payable without controlling for the geographical location of the property, i.e.:

$$p_{idbt} = \beta \tau_{dbt} + \delta_{bt} + \zeta' x_{idbt} + \varepsilon_{idbt} \tag{3.1}$$

where p_{idbt} is the price of house i in Borough d, band b at time t; τ_{dbt} is the council tax amount for a house in Borough d, band b at time t; δ_{bt} are year-band fixed effects; and x_{idbt} are controls which include the property size measured in squared meters, number of rooms, property type, age, duration and month which controls for seasonality in the housing market (Ngai and Tenreyro, 2014). Table C1 reports the results of regression (3.1); the first column provides the baseline result where month and year-band fixed effects are included in order to remove the mechanical correlation between increasing property prices and taxes over time and the fact that moving from band A to band H goes hand in hand with higher house prices. If we took this evidence at face value, we would conclude that the incidence of council tax is extremely large and statistically significant with a point estimate of -231.2. To give intuition, using a discount rate of $r = 4\%$ (similar to the risk-free rate observed in sample) this would roughly imply that an extra £1 in present value of taxes would lead to a drop in prices of $r \times \beta = 4\% \times 231.2 = £9.25$. It is obvious that this figure is only the artefact of the negative correlation between the value of properties and the average tax within councils as observed in Figure C11. Extremely negative coefficients are obtained in columns (2), (3) and (4) where we control for the property size, number of rooms, property type, whether the property is newly-built and whether it is a leasehold. The smallest of these coefficients in absolute value, i.e., -228.7 in column (3), would imply an incidence of $r \times \beta = 4\% \times 228.7 = £9.15$ which is still unreasonably high. Table C2 shows similar estimates when we include all the variables available as controls. To further corroborate the negative correlation between property prices and council taxes due to geographical selection, we provide the results of the following two-step estimation. First, we regress house prices on characteristics to obtain hedonic residuals:

$$p_{idbt} = \zeta' x_{idbt} + \varepsilon_{idbt} \tag{3.2}$$

For each Borough, band, year, we compute the median residual price ε^{med}_{dbt} and proceed to regress it on council tax amount payable including year-band fixed effects:

$$\varepsilon^{med}_{dbt} = \beta\tau_{dbt} + \delta_{bt} + \eta_{dbt} \tag{3.3}$$

The results are reported in Table C3. The vector of predictors x_{idbt} in the first-stage hedonic regression includes: month fixed effects in column (1); month, property size, number of rooms in column (2); month, property size, number of rooms and property type in column (3); and month, property size, number of rooms, property type and indicators for whether the property is newly-built and a leashold in column (4). Similarly, Table C4 reports results when the dependent variable in the second stage is the average hedonic residual $\bar{\varepsilon}_{dbt}$ per Borough, band, year, i.e.:

$$\bar{\varepsilon}_{dbt} = \beta\tau_{dbt} + \delta_{bt} + \eta_{dbt} \tag{3.4}$$

Both tables confirm the previous finding that Boroughs with higher house values tend to impose lower council tax bills: the coefficients are negative and statistically significant, ranging from -183.6 to -368.4.

The results provided so far imply that special care needs to be taken before using the geographical variation in council taxes to estimate their incidence on house prices. For this reason in our identification strategy we compare only houses that lie extremely close, i.e., no more than 500 meters and mainly closer than 200 meters, to the border between two adjacent Boroughs in order to disentangle the actual incidence of the tax from the geographical distribution of taxes across Boroughs. Throughout the rest of the paper, the reader should bear in mind that the geographical distribution of council taxes entails that any estimated incidence is, at most, an upper bound for the *true* incidence. This is because, if buyers value certain characteristics upon purchasing a house, these should be capitalised in the house price which, in this case, acts almost like a sufficient statistic for their value; the results of Figure C11 and Tables C1-C4 signal that houses with more highly valued characteristics (and higher prices) tend to be located in Boroughs with lower taxes, thus inflating any estimate of tax incidence. A second and more subtle reason why we can only estimate an upper bound for the incidence has to do with our identification strategy. By comparing similar dwellings on opposite sides of a border, we implicitly assume that the buyer always has an outside option during the price bargaining process. As a result, the buyer would be much more elastic than an otherwise identical buyer involved in the purchase of a house located in the heart of a Borough where there is no outside option in terms of council tax. We show in Section 3.4 that the seller bears the full incidence of the tax at the border, while that is not necessarily the case at an interior point. In general, even in the absence of perfect substitutes across council borders, it is reasonable to conjecture that the incidence is still much larger at the border compared to the council centre, where the agent would have to move long distance in order to pay a different council tax rate.

In the next section we describe the identification strategies that allow us to estimate the incidence of council taxes as precisely as possible given the present setting, bearing in mind that any attempt is likely to result in an over-estimation of the *true* incidence.

3.2.2 Identification Strategies

We use two different identification strategies to measure an upper bound of the incidence of council tax on property prices: regressions grids and a matching algorithm.

Regression Grids

The first strategy compares houses that lie in close proximity by dividing the area of London in a grid and assigning a fixed effect to each square in the grid. By doing so, we are de-facto comparing two houses that are otherwise identical but lie on opposite sides of a given border between two Boroughs. Figure C12 graphically depicts our first approach. Panel C12a shows a grid of squares with equal sizes superposed on a map of London. Panel C12b shows a more detailed picture of the Boroughs in the centre[10]. We then proceed to select the squares that have two houses that: are sold in the same year, are in the same council tax band and lie on opposite sides of the border; Panel C12b displays in blue examples of such squares. It can be noticed that we discard observations for which the border is located on the Thames River bank. To avoid relying on an arbitrary division, we use three different grids, namely one grid divides the area in 50×50 squares, another divides it in 100×100 squares and, finally, the last grid is a 150×150 one. These squares have an approximate size of 800 meters, 400 meters and 250 meters, respectively. While the maximal possible distance between houses can be inferred as $\sqrt{2} \times$ square side length, we choose to remove observations that are more than 500 meters far from the border. Figure C13 shows the distribution of distances to the border for our different specifications. As mentioned, no house lies more than 500 meters away from the border, and most of the observations are about 200 meters away from the closest border. As we proceed to refine our grids by subdividing into a larger number of squares, we can see that we lose observations in the 200 meters-500 meters range; this reduces our power significantly, but ensures that we compare houses that are indeed in very close proximity.

Our strategy consists of running within square regressions whereby we compare houses that are sold in the same year and in the same council tax band, specifically:

$$p_{ibgdt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibgdt} + \varepsilon_{ibgdt} \tag{3.5}$$

[10] The three main Boroughs depicted in the picture are, starting from left, Hammersmith and Fulham, Kensington and Chelsea and the City of Westminster.

where p_{ibgdt} is the price of house i, in council tax band b, grid square g, Borough d, and year t; τ_{bdt} is the council tax amount for band b, Borough d in year t; and x_{ibgdt} are house-specific controls. The presence of the band-grid square-year fixed effects δ_{bgt} guarantees that the regression compares houses that are in the same square, same council tax band and are sold in the same year, implying that our identification assumption is that they systematically differ only due to the amount of council tax paid, after partialling out the effect of house characteristics that we add to increase our precision. It should be noticed that, as mentioned above, *better* Boroughs, i.e., Boroughs with higher average prices, tend to have lower council taxes, implying that - if we leave some hidden characteristic out of our regression - the estimate of β is most likely going to overstate the *true* incidence. To give an example, while highly unlikely given the sharp nature of the borders, one could argue that there is a name tag value of living in certain Boroughs over others, for instance, a house in Westminster commands a premium over a similar house on the other side of the border in Brent. The fact that Westminster has a lower tax compared to Brent implies that we would overestimate the incidence of the tax because of the name tag value of living in Westminster. In general, to reverse this bias and claim that the *true* incidence might be higher than the one we estimate, the reader should think of some hidden characteristic that systematically causes people to prefer living in a Borough with worse amenities compared to a Borough with better ones.

The following section presents our second identification strategy which relies on a matching estimator rather than grid squares fixed effects.

Matching Estimator

Our second identification approach consists of pairwise matching of houses on opposite sides of a given border. To find the closest match, we need to define a distance: in what follows, we rely on a Euclidean distance and a distance based on a linear model. Under the first one, we restrict the possible matches to be: no more than 500 meters away from each other, sold in the same year, in the same council tax band, and to both be either old or newly-built and freehold or leasehold properties. For each property we then choose the closest match as the one minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$, where i is the original property, j indexes the possible matches on the other side of the border, x_{ik} are house i characteristics and x_{jk} are house j characteristics. We then run within-pair regressions:

$$p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt} \tag{3.6}$$

where δ_{ij} are ij-pair dummies and x_{ibdt} are house i-specific features. The second choice of distance is based on a linear pricing model:

$$p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \tag{3.7}$$

where x_{it} similarly contains house-specific characteristics as above. We then compute the model-predicted price $\hat{p}_{it} = \hat{\alpha} + \hat{\beta}' x_{it}$. As before, we restrict the pairing to houses sold in the same year, band, old/new and leasehold/freehold categories and no further than 500 meters from each other. For each property i we pick the closest match j as the one that minimises the following distance: $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}|$. To estimate the incidence, we run within pair-regressions as in equation (3.6) where the δ_{ij} dummies are determined according to the new matching algorithm. As in Section 3.2.2 the identification is valid as long as the only systematic difference within pairs is the amount of council tax. As previously explained, any other omitted variable would most likely lead us to estimate an upper bound for the incidence, given the geographical distribution of council taxes.

3.3 Results

3.3.1 Grid Estimator

Table C5 presents the results of the grid regressions described in Section 3.2.2 where we use a 50×50 grid and include band-grid square ID-year fixed effects to compare the effect of council taxes on properties in the same band, sold in the same year, located in the same grid square but on opposite sides of a border as in equation (3.5).

The controls we include are as follows: column (1) uses month fixed effects to control for housing market seasonality; column (2) adds number of rooms fixed effects and controls for property size; column (3) also adds property type fixed effects, and; column (4) includes an indicator for newly-built and leasehold properties. These are our default specifications throughout the rest of the paper. In all columns the coefficient on council taxes is statistically indistinguishable from zero and always with the wrong sign. The lack of significance cannot be attributed to lack of statistical power in the regressions given that other control variables are always strongly statistically significant. For instance, the effect of one additional squared metre ranges between £4,537 and £4,627, newly-built properties command a premium of about £33,400 and freehold properties sell for £76,000 more relative to leaseholds. The same conclusion can be drawn from Table C6 where we expand the regressions to include all available house price predictors, showing that even relatively minor characteristics such as the number of lighting outlets or the presence of fireplaces in the property have a significant effect on prices.

Table C7 displays the grid regression results for grids of different sizes: column (1) uses a grid that divides the London area into 50×50 squares, column (2) 100×100, and column (3) 150×150. This might help to alleviate concerns that grids made of large squares might be comparing houses that are rather distant from each other. The specification is otherwise same as the one in column (4) of Table C5. The coefficient on council tax remains statistically insignificant and the point estimate varies from positive to negative across columns: this is precisely what we should expect when a regressor has no effect on the outcome variable and simply reacts to the noise in the sample. The fact that the R-squared is very high (between 77% and 83%) and that all other coefficients are precisely estimated confirms our previous finding that the incidence of the council tax is indistinguishable from zero. In Table C8 we augment the regressions by adding all additional house characteristics: the coefficient on council tax ranges from -11.8 to 75.4 and is never statistically lower than zero.

To make sure that the confounding effect of the stamp duty notches does not play a role in our estimation results, Table C9 presents the results of the grid regressions when we remove the two main stamp duty notches at £250,000 and £500,000. Column (1) excludes only the first notch, column (2) the second, and column (3) removes both. The results are virtually unchanged, with the incidence still being statistically insignificant, small in magnitude, and always displaying the wrong sign. As previously mentioned, the large R-squared and the fact that the remaining coefficients are precisely estimated guarantees that this is not due to lack of power.

Finally, Tables C10 and C11 provide estimates of council tax incidence using a similar two-step approach as in Tables C3 and C4, i.e., by first obtaining residual hedonic prices as follows:

$$p_{ibdgt} = \zeta' x_{ibdgt} + \varepsilon_{ibdgt} \tag{3.8}$$

and subsequently regressing the median or average hedonic residuals for each Borough, band, grid square and year on council tax amounts:

$$\varepsilon^{med}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt} \tag{3.9}$$

$$\bar{\varepsilon}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt} \tag{3.10}$$

where δ_{bgt} are band-grid square-year fixed effects included to ensure that we compare values of houses in the same council tax band, sold in the same year and located in the same square of the grid. As usual, we restrict the analysis to grid squares with at least two houses located on different sides of a border and present the four standard specifications. The results confirm the previous finding: both the median and average hedonic residuals are not decreasing in the council tax amount paid, suggesting that the incidence of this tax on house prices is not different from zero.

In the following section we supplement the evidence by presenting results using our second identification strategy.

3.3.2 Matching Estimator

Tables C12, C13 and C14 show the results of our second estimation approach where we explicitly match similar dwellings on opposite sides of a border as described in Section 3.2.2. As previously mentioned, all the results are obtained using housing pairs on opposite sides of a border no more than 500 metres apart, sold in the same year, in the same council tax band and which are both either old or newly-built and leasehold or freehold properties. Table C12 displays the results where closest pairs have been determined by minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$, where the vectors x_i and x_j consist of property size and number of rooms in columns (1) and (2), and also energy cost in columns (3) and (4). All the variables are standardised to be comparable. This procedure leads to 57,612 and 57,323 observations of property pairs with 71,578 and 71,656 unique transactions in columns (1)-(2) and (3)-(4), respectively[11]. After having obtained the pairs, we run the regression specified in equation (3.6). The presence of δ_{ij} pair fixed effects amounts to regressing the difference in prices of matched houses on the difference in council tax paid, controlling for other property characteristics along which the matched properties may differ. Consistent with the results obtained with the grid estimator, none of the coefficients on council tax is statistically significantly negative. As pointed out before, this result is not attributable to lack of statistical power: for instance, the coefficient on size is highly statistically significant and has the same order of magnitude as the ones obtained with the earlier estimator[12]. Table C13 confirms these findings under the linear matching algorithm where pairs are chosen by minimising the distance $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}|$, where the predicted prices $\hat{p}_{it}$ and $\hat{p}_{jt}$ are obtained from a linear model as in equation (3.7). As before, columns (1) and (2) match properties based on size and number of rooms, while columns (3) and (4) add energy cost. Finally, Table C14 presents the last set of results for the linear model where we allow each property to be paired with

[11] Notice that any given transaction can be the closest match for more than one property. In order to take care of this redundancy we cluster standard errors at the transaction ID level.

[12] Notice that, compared to the default specifications used in Tables C1, C5, C7 and C9, the indicators for newly-built and leasehold properties have been dropped given that properties are constrained to be identical along these dimensions.

more than one similar property on the other side of the border, as long as the absolute difference in predicted prices is less than 30% of the largest predicted price, namely: $|\hat{p}_{it} - \hat{p}_{jt}| < 0.3 \times \max\{\hat{p}_{it}, \hat{p}_{jt}\}$. While the point estimates range between -5.24 and -8.19, none of the coefficients is statistically different from zero as in all previous specifications. We shed more light on the interpretation of these and the previous results in Section 3.4.1.

The empirical findings above demonstrate that council tax differences never significantly explain house price differences. Moreover, while absence of evidence, namely the fact that agents seem to be insensitive to taxes that are postponed to the future, does not directly imply evidence of absence, many point estimates are positive and hence with the wrong sign. Bearing these estimates in mind, in the next section we develop a simple model that allows us to propose a plausible explanation for the above results. We subsequently calibrate the model using a Bayesian approach informed by all of the above estimates and briefly discuss policy implications.

3.4 Model

In what follows, we present a simple multi-period model of housing-consumption choice in order to calibrate the above results. We begin with the optimisation problem of an agent who chooses at time $t = 0$ an infinite stream of consumption $\{c_t\}_{t=0}^{\infty}$ and a composite housing good h:

$$\max_{\{\{c_t, d_t\}_{t=0}^{\infty}, h, \mathbb{1}_{\{A\}}, \mathbb{1}_{\{B\}}\}} \quad U(\{c_t\}_{t=0}^{\infty}, h) = c_0 + \sum_{t=1}^{\infty} \beta^t u(c_t) + \sum_{t=0}^{\infty} \beta^t \log(h) \tag{3.11}$$

$$s.t. \quad c_0 + h(p_{A0}\mathbb{1}_{\{A\}} + p_{B0}\mathbb{1}_{\{B\}} + \tau_S) \leq w_0 + d_0 \tag{3.12}$$

$$c_t + h(\tau_{At}\mathbb{1}_{\{A\}} + \tau_{Bt}\mathbb{1}_{\{B\}}) + d_{t-1}(1+r) \leq w_t + d_t \quad t = 1, 2, 3, ... \tag{3.13}$$

$$d_t \leq \alpha h(p_{At}\mathbb{1}_{\{A\}} + p_{Bt}\mathbb{1}_{\{B\}}) \quad t = 0, 1, 2... \tag{3.14}$$

For simplicity, the utility of the agent is chosen to be time-separable and separable in consumption and housing. The utility function is quasi-linear in c_0 in order to get rid of income effects, as is standard practice in the public finance literature. For tractability and to separate the effects of stamp duty and council tax, the agent purchases the housing good only once at $t = 0$. There are two Boroughs, A and B, with exogenously chosen and potentially different council tax rates. We assume that there is equal supply of housing in both Boroughs[13]. Equation (3.12) is the first-period budget constraint: the agent spends his initial endowment w_0 on consumption c_0 and the after-tax cost of his housing demand h. When he buys a house, the agent pays the pre-tax price $p_{i0}, i = A, B$, and, in addition, he also needs to pay the stamp duty tax τ_S hereby assumed to be proportional to the quality-adjusted level of housing demand. If his total demand exceeds his initial endowment, the agent can borrow additional funds d_0 for one period at the risk-free rate. The budget constraints for all subsequent periods are identical and given by equation (3.13): from time $t = 1$ onwards, the agent spends his endowment w_t on his optimal consumption choice c_t and to pay the council tax $\tau_{it}, i = A, B$, that corresponds to the Borough where he has chosen to locate at time $t = 0$. He also needs to repay his short-term debt from the previous period inclusive of interest $d_{t-1}(1 + r)$, and is allowed to borrow again at the same terms in order to balance his budget constraint. Finally, the last constraint in equation (3.14) is the financing constraint: the agent cannot borrow more than a fraction α of the pre-tax cost of his housing demand. This can potentially generate very large incidence for the stamp duty tax since the lump sum nature of this tax tightens the leverage constraint. The Lagrangian for the above problem can be written as:

$$
\begin{aligned}
\mathcal{L} =& U(\{c_t\}_{t=0}^{\infty}, h) - \lambda_0(c_0 + h(p_{B0} + \tau_S) - w_0 - d_0) \\
& - \sum_{t=1}^{\infty} \lambda_t(c_t + h\tau_{Bt} + d_{t-1}(1 + r) - w_t - d_t) - \sum_{t=0}^{\infty} \mu_t(d_t - \alpha h p_{Bt}) \\
& - h\mathbb{1}_{\{A\}} \left[\lambda_0(p_{A0} - p_{B0}) + \sum_{t=1}^{\infty} \lambda_t(\tau_{At} - \tau_{Bt}) - \alpha \sum_{t=0}^{\infty} \mu_t(p_{At} - p_{Bt}) \right]
\end{aligned}
\tag{3.15}
$$

[13]This assumption is crucial and de-facto eliminates the potential for a differential elasticity of supply with respect to council taxes at the border. We consider this assumption quite reasonable given that the greatest majority of the housing stock in London has been constructed well before the introduction of this tax in the early 90s as shown in Figures C8 and C9.

where we use the fact that $\mathbb{1}_{\{B\}} = 1 - \mathbb{1}_{\{A\}}$. Notice that the Lagrangian is monotone in the choice of Borough $\mathbb{1}_{\{A\}}$, therefore, the choice of where to locate can be separated from the consumption and housing-quality choices. The agent chooses to live in Borough A if:

$$p_{A0} - p_{B0} \leq -\sum_{t=1}^{\infty} \frac{\lambda_t}{\lambda_0}(\tau_{At} - \tau_{Bt}) + \alpha \sum_{t=0}^{\infty} \frac{\mu_t}{\lambda_0}(p_{At} - p_{Bt}) \tag{3.16}$$

i.e., if the price differential between the same-quality house in Boroughs A and B more than compensates for the present value of the difference in future council tax payments and the collateral value of the house. In equilibrium, markets clear if equation (3.16) holds with equality which, from now onwards, we assume to be the case. Assuming that the agent is indifferent between living in Boroughs A and B, we proceed by suppressing the Borough subscripts and denote the price of the house as p_t and the council tax as τ_t. The first-order conditions for an interior solution are:

$$1 = \lambda_0 \tag{3.17}$$

$$\beta^t u'(c_t) = \lambda_t \quad \forall t = 1, 2, 3, ... \tag{3.18}$$

$$-\lambda_t + \lambda_{t+1}(1 + r) + \mu_t = 0 \quad \forall t = 0, 1, 2, ... \tag{3.19}$$

$$\frac{h^{-1}}{(1-\beta)} = \lambda_0(p_0 - \alpha\frac{\mu_0}{\lambda_0}p_0 + \tau_S) + \sum_{t=0}^{\infty} \lambda_{t+1}\tau_{t+1} - \sum_{t=0}^{\infty} \lambda_{t+2}\frac{\mu_{t+1}}{\lambda_{t+2}}\alpha p_{t+1} \tag{3.20}$$

Combining the first-order conditions for consumption and for the optimal debt choice, we obtain the following Euler equation:

$$\frac{\lambda_{t+1}}{\lambda_t} = \beta\frac{u'(c_{t+1})}{u'(c_t)} = \frac{1}{1 + r + \frac{\mu_t}{\lambda_{t+1}}} \tag{3.21}$$

The above Euler equation implies that the agent's discount factor is equal to the inverse of the risk-free rate and a liquidity premium $\frac{\mu_t}{\lambda_{t+1}}$, arising from the fact that the house has some collateral value. In order to simplify the exposition, we assume that in equilibrium the liquidity premium is constant and equal to $\frac{\mu_t}{\lambda_{t+1}} = k$, that house prices grow at a constant rate g, i.e., $p_{it} = p_{i0}(1+g)^t$, and council tax amounts grow at a constant rate $\tilde{g}$, i.e., $\tau_{it} = \tau_{i1}(1+\tilde{g})^{t-1}$. Re-arranging equations (3.16), (3.20) and (3.21), we obtain the final no-arbitrage condition and housing demand:

$$(p_{A0} - p_{B0})\left(1 - \frac{\alpha k}{r+k-g}\right) = -(\tau_{A1} - \tau_{B1})\frac{1}{r+k-\tilde{g}} \tag{3.22}$$

$$\frac{h^{-1}}{(1-\beta)} = p_0\left(1 - \frac{\alpha k}{r+k-g}\right) + \tau_S + \frac{\tau_1}{r+k-\tilde{g}} \tag{3.23}$$

The first equation is the equilibrium condition of how house prices should behave across Boroughs: the house price differential, after having taken into account the collateral value $\frac{\alpha k}{r+k-g}$, needs to match (the negative of) the present value of the council tax differential. The second equation states that the agent's marginal utility of housing is equal to the house price inclusive of (the present value of) all taxes and collateral value. It is important to note that the no-arbitrage condition (3.22) in general gives a different incidence compared to the one obtained from the housing demand (3.23). This is because the former holds only at the border between two Boroughs where the outside option, i.e., the option to buy an otherwise identical house on the other side of the border, implies that the supply bears the whole burden of the tax. In particular, from equation (3.22) we obtain an incidence of:

$$\frac{dp_0}{d\tau_1} = -\frac{1}{r+k-\tilde{g}} \times \frac{r+k-g}{r+(1-\alpha)k-g} \tag{3.24}$$

On the other hand, for both houses on the border as well as houses in the middle of a given Borough we can define the optimal demand from equation (3.23) as $D(p_0, \tau_1, \tau_S) = h^*(p_0, \tau_1, \tau_S)$. Equating with the optimal supply, $S(p_0) = D(p_0, \tau_1, \tau_S)$, and after total differentiation we obtain the standard formula for the incidence:

$$\frac{dp_0}{d\tau_1} = -\frac{\frac{\partial D}{\partial \tau_1}}{\frac{\partial D}{\partial p_0} - \frac{\partial S}{\partial p_0}} = -\frac{1}{r+k-\tilde{g}} \times \frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \tag{3.25}$$

where $\tilde{\eta}_S = \frac{\partial S}{\partial p_0}\frac{p_0}{S}\frac{p_0\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau_1}{r+k-\tilde{g}}}{p_0} = \eta_S\frac{p_0\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau_1}{r+k-\tilde{g}}}{p_0}$ is a slightly modified version of the supply elasticity η_S that takes into account the price inclusive of taxes and collateral value. In general, we have that:

$$\frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \leq \frac{r+k-g}{r+(1-\alpha)k-g} \tag{3.26}$$

implying that the incidence at the border between Boroughs is an upper bound for the *true* council tax incidence as long as the modified elasticity of supply is non-negative, i.e., $\tilde{\eta}_S \geq 0$. Notice that the modified elasticity of supply $\tilde{\eta}_S$ is positive as long as the true elasticity of supply η_S is positive.

3.4.1 Calibration

The model in the previous section allows us to better interpret the empirical results of Section 3.3. By using equations (3.22), (3.23) and (3.24) we get[14]:

$$\frac{dp_0}{d\tau_1} = \frac{dp_0}{d\tau_S} \times \frac{1}{r+k-\tilde{g}} \tag{3.27}$$

i.e., the incidence of the council tax can be interpreted as the present value of the sum of the incidence of the stamp duty tax discounted at the liquidity-adjusted cost of capital $r+k$ with growth rate $\tilde{g}$. In what follows we use the results in Tables C5 - C14 and provide further direction on how to interpret them. We treat each estimate as a separate model m. Conditional on the model being true and given a common prior distribution $p(\beta_\tau|m) = p(\beta_\tau)$ about the true incidence of council tax and the likelihood function of the data $p(y|\beta_\tau, m)$ we can use Bayes' rule to express the

[14]This assumes that $\tilde{\eta}_S = 0$, i.e., that the supply of housing is fixed in the short term.

posterior distribution for the incidence under each model m as:

$$p(\beta_\tau|y,m) = \frac{p(y|\beta_\tau,m) \times p(\beta_\tau)}{\int p(y|\beta_\tau,m) \times p(\beta_\tau) d\beta_\tau} \tag{3.28}$$

We then proceed to obtain the model-averaged posterior distribution as:

$$p(\beta_\tau|y) = \sum_m p(\beta_\tau|y,m)p(m|y) \tag{3.29}$$

The computational burden of equation (3.29) is significant, therefore, we proceed with the simplifying assumptions described in Appendix C.4. We always start from a normally-distributed prior $\beta_\tau \sim \mathcal{N}(b_\tau, \sigma_\tau^2)$ and likelihood function which leads to a normal posterior. As detailed in Appendix C.4 the mean of the prior is chosen by calibrating the parameters g, $\tilde{g}$, r and α based on historical data and matching the stamp duty incidence to results in Best and Kleven (2018). For robustness we also vary the precision of the prior and provide results for five different specifications: $p(\beta_\tau) = \mathcal{N}(-150, 50^2), \mathcal{N}(-100, 50^2), \mathcal{N}(-50, 50^2), \mathcal{N}(-150, 75^2), \mathcal{N}(-50, 25^2)$.

Figure C15 plots the model-averaged density of the posterior distribution for the council tax incidence. Panel (a) displays the posterior density for a constant standard deviation of the prior of 50, while (b) for a standard deviation equal to half the prior mean. It can be noted that the shape of the posterior is similar across specifications and that it displays a significant shift of mass toward zero. Table C15 provides the quantiles, the mode and the mean of the posterior distribution of the incidence. The median posterior incidence ranges between -22.87 and -2.17, well below the median implied by the model calibration which has informed the prior. The last column reports the ratio between the two, giving the implied attenuation bias displayed by agents. Given the model parameters the price reaction to council taxes is between 4% and 37% of what the price reaction to the stamp duty tax would imply from agents who fully perceive the tax.

The results above become striking once coupled with the extent to which house buyers react to stamp duty taxes. When buyers are liquidity-constrained, their effective discount rates become large and, therefore, one might be tempted to attribute the previous evidence solely to extreme discounting of future cash flows. If we are willing to take this view, we would have to assume discount rates ranging between 23.4% and 231.9% in order to fit the posterior estimates of the council tax incidence. Moreover, it should be noted that every estimate of the council tax incidence is conditioned on an estimate of the stamp duty incidence, i.e., the discount rate is not a free parameter in the calibration. To put it differently, changing the discount rate to match a reasonable incidence for the council tax would lead to an incidence of the stamp duty tax that is inconsistent with current estimates in the literature. The fact that the incidence of the stamp duty is large but not extreme implies that the liquidity premium cannot be the only source of the low council tax incidence. Third, in our estimation we use relatively concentrated priors around the model-informed incidence; had we allowed the likelihood to dominate by assigning diffuse priors, we would have obtained much lower estimates compared to the conservative ones provided so far. One way to explain these findings is by hypothesising that, when buying their properties, agents discount tax payments that happen in the future disproportionately compared to those that occur concurrently with the purchase. It is difficult to argue that this might be due to uncertainty associated with council tax payments given that differences in council tax amounts across Boroughs are very smooth and predictable as shown in Figure C2. This leaves us with another plausible alternative explanation: agents fail to fully internalise the difference in council tax payments across Boroughs upon purchasing a property, either because this is much less salient compared to the stamp duty tax[15], or because they fail to appreciate the magnitude of its present value[16]. Notice also that the results so far suggest that there is somebody who does not take the council tax differentials into account in a fully-rational way, but this does not need to be the house buyer: our previous analysis goes through even if the buyer is fully aware

[15] It is also possible that the tax is fully salient to agents but, due to mental accounting, they fail to integrate its present value into the house price they are willing to pay. Other explanations could be related to search costs and cognitive costs.

[16] For a property in band D worth, say, £300,000, the stamp duty tax in 2018 would amount to £9,000. If the buyer could choose whether to buy the property in the Borough of Camden or the Borough of Westminster, the difference in council tax would amount to about £778 in 2018 which, in present value using a discount rate of 4%, would be equal to £19,450, more than twice the value of the stamp duty tax.

of the tax and hopes to shift its incidence onto the subsequent buyer, or the renter in the case of buy-to-let property transactions[17].

Motivated by these findings, we explore some policy implications in the following section.

3.4.2 Implications for Tax Policy

Given the results in the previous section, it seems reasonable to argue that agents fail to fully perceive deferred taxes. As a result, we propose a modified version of the model above that allows for non-fully salient taxes. We extend our analysis to properties that are potentially far from the border and, therefore, allow the elasticity of supply η_S to be non-zero. Recall that the incidence estimates coming from the border in Section 3.3 are an upper bound for the incidence in the middle of Boroughs. For simplicity, let us assume we are in an equilibrium where the leverage constraint (3.14) is binding, i.e., $d_t = \alpha h p_t$. If we multiply each of the constraints (3.12) and (3.13) by $\frac{1}{(1+r+k)^t}$ and add them together, we obtain the following consolidated budget constraint:

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \qquad (3.30)$$

where $\tilde{p} = p_0 \left(1 - \frac{\alpha k}{r+k-g}\right) + \tau_S + \frac{\tau_1}{r+k-\tilde{g}}$ is the tax-inclusive house price. For simplicity of exposition, define $p = p_0 \left(1 - \frac{\alpha k}{r+k-g}\right)$ and $\tau = \frac{\tau_1}{r+k-\tilde{g}}$, so that we can rewrite $\tilde{p} = p + \tau_S + \tau$. Following Chetty et al. (2009), Farhi and Gabaix (2020) and Goldin (2015), we assume that the agent misperceives taxes with attenuation factor γ, i.e., he solves the following maximisation problem:

$$\max_{\{\{c_t\}_{t=0}^{\infty}, h\}} U(\{c_t\}_{t=0}^{\infty}, h) = c_0 + \log(h) + \sum_{t=1}^{\infty} \beta^t \left(u(c_t) + \log(h)\right) \qquad (3.31)$$

[17]Note that we largely interpret the results as evidence of overpricing. Another possibility is that the properties on the low council tax side of borders are relatively underpriced and it is, therefore, sellers who fail to incorporate the tax discount into their ask price.

s.t.

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}_\gamma h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \tag{3.32}$$

where the perceived house price is:

$$\tilde{p}_\gamma = p + \tau_S + \gamma\tau, \quad \gamma \in [0,1] \tag{3.33}$$

Recall from the previous section that the attenuation factor for the council tax implied by the data ranges between 0.04 and 0.37. Notice that while the agent perceives the above budget constraint, he has to satisfy the actual budget constraint (3.30) given by the rational model. As pointed out in Reck (2016), it is crucial to decide what choice variable bears the burden of adjustment. Given our assumption about the quasi-linear utility function in first-period consumption c_0, it is natural to let c_0 be the shock absorber. This choice amounts to assuming the following train of events: 1) the agent misperceives the council tax he will have to pay going forward and, as a result, buys "too much" quality-adjusted housing; 2) following this, he realises that the actual amount of taxes he will have to pay is beyond his budget; 3) consequently, the agent adjusts his consumption in the first period keeping everything else constant. Denoting the observed demands as $\hat{c_0}, \hat{c_t}, \hat{h}$, and the optimal demands absent any behavioural frictions as c_0^*, c_t^*, h^*, we have the following first-order conditions:

$$\hat{c_t} = [u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right) = c_t^* \tag{3.34}$$

$$\hat{h} = [(1-\beta)\tilde{p}_\gamma]^{-1} \neq [(1-\beta)\tilde{p}]^{-1} = h^* \tag{3.35}$$

$$\hat{c_0} = I - \sum_{t=1}^{\infty} \frac{\hat{c_t}}{(1+r+k)^t} - \hat{h}\tilde{p} \neq c_0^* \tag{3.36}$$

As previously mentioned, the optimality condition for future consumption remains as before. However, equation (3.35) shows that the agent demands "too much" housing due to the fact that the perceived price $\tilde{p}_\gamma$ is lower than the true price $\tilde{p}$, as long as $\gamma < 1$. As a result, because of quasi-linearity in the utility function,

$\hat{c}_0$ adjusts to absorb the reduction in available income. The previous discussion highlights the fact that misperception of the house price affects both consumption and housing demand, albeit in opposite directions. This implies that a benevolent social planner needs to carefully balance the two distortions when setting the optimal tax policy. To see this more formally, let us adopt the approach of Goldin (2015) and assume that the Government chooses the optimal (property) tax combination in order to raise a fixed amount of revenue and maximise the utility of the buyer[18]. For convenience, define the present value of council tax revenue from the Government's point of view, discounted at the risk-free rate, as $\tilde{\tau} = \frac{\tau_1}{r-\tilde{g}}$. The total revenue raised from a given buyer is:

$$R = (\tau_S + \tilde{\tau})h = \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)h \tag{3.37}$$

The second equality of the above equation shows that the Government discounts the revenue raised through council taxes at a lower rate than agents due to the presence of borrowing constraints. The Government can twick the two taxes to maintain revenue-neutrality. In particular, a revenue-neutral tax change is such that:

$$\left[h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau_S}\right]\Delta\tau_S = -\left[\frac{r+k-\tilde{g}}{r-\tilde{g}}h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau}\right]\Delta\tau \tag{3.38}$$

This implies that the change in stamp duty per unit change in council tax needed to maintain revenue-neutrality is:

$$\frac{\Delta\tau_S}{\Delta\tau} = -\frac{\frac{r+k-\tilde{g}}{r-\tilde{g}}h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau}}{h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau_S}} = -\frac{\frac{r+k-\tilde{g}}{r-\tilde{g}}h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\theta_\tau\frac{\partial h}{\partial p}}{h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\theta_{\tau_S}\frac{\partial h}{\partial p}} \tag{3.39}$$

where $\theta_{\tau_S} = \frac{\frac{\partial h}{\partial \tau_S}}{\frac{\partial h}{\partial p}}$ and $\theta_\tau = \frac{\frac{\partial h}{\partial \tau}}{\frac{\partial h}{\partial p}}$ tell us how responsive the demand is with respect to taxes relative to pre-tax prices. From equations (3.33) and (3.35) we infer that $\theta_{\tau_S} = 1$ and $\theta_\tau = \gamma$ in our model. The indirect utility function for an inattentive

[18]In what follows, we abstract from analysing the effect on the utility of the seller.

agent is:

$$V(p,\tau_S,\tau) = I - \sum_{t=1}^{\infty} \frac{\hat{c}_t}{(1+r+k)^t} - \hat{h}(p+\tau_S+\tau) + \sum_{t=1}^{\infty} \beta^t u(\hat{c}_t) + \frac{\log(\hat{h})}{(1-\beta)} \quad (3.40)$$

where $\hat{c}_t = [u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right)$ and $\hat{h} = \hat{h}(p,\tau_S,\tau) = [(1-\beta)(p+\tau_S+\gamma\tau)]^{-1}$ from the agent's first-order conditions. Differentiate the indirect utility function above to obtain:

$$\frac{dV}{d\tau} = -\hat{h}\left(\frac{dp}{d\tau} + \frac{\partial \tau_S}{\partial \tau} + 1\right) + \left[\frac{\partial U}{\partial h} - (p+\tau_S+\tau)\right]\left[\frac{dp}{d\tau} + \theta_{\tau_S}\frac{\partial \tau_S}{\partial \tau} + \theta_\tau\right]\frac{\partial \hat{h}}{\partial p} \quad (3.41)$$

where $\frac{dp}{d\tau} = \frac{\partial p}{\partial \tau} + \frac{\partial p}{\partial \tau_S}\frac{\partial \tau_S}{\partial \tau}$ is the total incidence of the council tax after having taken into account the shift in stamp duty to guarantee revenue neutrality. As in Goldin (2015), the change in welfare can be decomposed into four components: the first part, i.e., $-\hat{h}\left(\frac{dp}{d\tau} + \frac{\partial \tau_S}{\partial \tau} + 1\right)$ measures the direct welfare effect of a tax shift due to the alleviation of the borrowing constraint; the second part, i.e., $\left[\frac{\partial U}{\partial h} - (p+\tau_S+\tau)\right]$ is the behavioural wedge and it represents the difference between perceived and actual prices; the third component, i.e., $\left[\frac{dp}{d\tau} + \theta_{\tau_S}\frac{\partial \tau_S}{\partial \tau} + \theta_\tau\right]$ is equal to the change in prices as perceived by the agent; and the fourth component, i.e., $\frac{\partial \hat{h}}{\partial p}$ is the impact of a change in prices on demand for housing. With no bias, i.e., when $\gamma = 1$ the perceived price is equal to the actual price and the envelope theorem ensures that the second component above is equal to zero. As a consequence, the optimal tax policy depends on the sign of the first term[19]. If this is positive, it is optimal for the government to set $\tau_S = 0$, if negative, $\tau_S = R$. It is easy to show that when $\gamma = 1$ this term is unambiguously positive as long as $\eta_S > 0$. The Government should then choose a zero stamp duty tax in order to alleviate the agent's liquidity constraint. In the presence of biases, however, there is a trade-off between the two inefficiencies: 1) the liquidity constraint and differences in salience make increasing the stamp duty tax less efficient than raising the council tax; 2) on the other

[19]Notice that $\frac{\partial \tau_S}{\partial \tau} < -1$ because $r+k-\tilde{g} > r-\tilde{g}$, $\theta_\tau < \theta_{\tau_S}$ and $\frac{\partial h}{\partial p} < 0$. The above assumes that $\frac{r+k-\tilde{g}}{r-\tilde{g}}h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau} > 0$ and $h + \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau_S} > 0$, i.e., the Government is on the upward sloping part of the Laffer curve. The term $\frac{\partial p}{\partial \tau} + \frac{\partial p}{\partial \tau_S}\frac{\partial \tau_S}{\partial \tau}$ is usually positive since agents react less to a decrease in council tax relative to a revenue-neutral increase in the stamp duty.

hand, raising the council tax causes a shift in demand away from c_0 which in our example is the shock absorber. In the extreme case when there are no liquidity constraints, it is optimal to impose no council tax. Otherwise, the problem of the social planner amounts to choosing the optimal combination of stamp duty and council tax to jointly solve the following two equations:

$$\hat{h}\left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right) = R \tag{3.42}$$

$$\frac{dV}{d\tau} = 0 \tag{3.43}$$

Figure C16 reports the optimal mix of taxes computed for a house worth £430,000 which is the median value of properties in band D in 2017. The property pays a stamp duty of £11,500 and we assume that it pays a yearly council tax of £1,419.73, the in-sample median amount in the corresponding band and year. The upper panel shows how the optimal combination varies as a function of the discount rate $r+k$, while the bottom panel varies the attenuation parameter γ. The figures confirm the above intuition. From Figure C16a we can see that when the liquidity premium is zero, the optimal policy is to levy only the stamp duty tax. For a small liquidity premium there is an optimal mix that includes positive amounts of both taxes, however, the borrowing constraints become dominant fairly quickly and make it optimal to set a stamp duty of zero. Figure C16b, on the other hand, focuses on the effect of salience. Even when the council tax is entirely non-salient, i.e., $\gamma = 0$, it is still optimal to raise a little over 20% of revenue through it. As the tax becomes more salient, its distortionary effect on c_0 decreases, therefore, its proportion should increase, up to the point where it becomes the only form of taxation for γ greater than 0.25. It should be noted, however, that this assumes that tax policy changes do not affect any of the parameters. In practice, changing the tax mix can change the inattention parameter γ.

3.5 Conclusions

This paper studies the incidence of property taxes in the UK housing market. By using a geographical discontinuity approach, exploiting the considerable difference in council tax rates across London Boroughs, we show that agents significantly underreact to council taxes. Our empirical estimates of council tax incidence on house prices is never significantly negative and this lack of significance cannot be attributed to lack of power. This is in sharp contrast to the large stamp duty incidence estimated by Best and Kleven (2018) and suggests that agents do not pay sufficient attention to taxes deferred to the future, or possibly points to evidence of very large search frictions or other cognitive costs. In Section 3.4.2, we touch upon the policy implications of our findings, however, one should be aware of issues arising when manipulating tax rates given that there is no guarantee that changes in policies are not followed by changes in tax salience and therefore behaviour. The analysis in this paper relies on data from the residential property market, however, it can also be extended to other domains of tax policy. One general take-away from the present work is that transaction taxes, such as the stamp duty tax, have a large incidence on transaction prices while deferred taxes, such as the council tax, have a lower effect on prices but potentially higher impact on consumption choices. This implies that the optimal mix of taxes may be some combination of the two. The analysis can be extended, for instance, to financial securities where the fact that a transaction tax might be very distortionary does not imply that it is optimal to raise revenues only through capital gains[20] or dividend taxes.

The findings in the paper keep open the question of the nature of the channels through which inattentive households correct their mistakes and adjust their consumption policies, once neglected taxes materialise. Access to disaggregated expenditure data could help shed light on this matter: this can be done by analysing differences in consumption responses at the border between Boroughs, which we should expect to arise whenever agents fail to optimally account for tax differ-

[20] While the capital gains tax is a transaction tax, the fact that it is borne by the seller of the asset suggests that agents could still underreact to it as it is a deferred tax and, therefore less salient compared to a tax charged at the moment of purchase like the stamp duty tax.

ences and are forced to adjust their expenditures ex-post to meet their budget constraints.

Bibliography

Adam, Klaus, Dmitry Matveev, and Stefan Nagel, "Do Survey Expectations of Stock Returns Reflect Risk Adjustments?," *Journal of Monetary Economics*, 2021, *117*, 723–740.

Addoum, Jawad M. and Justin R. Murfin, "Equity Price Discovery with Informed Private Debt," *The Review of Financial Studies*, 2020, *33* (8), 3766–3803.

Adrian, Tobias, Paolo Colla, and Hyun Song Shin, "Which Financial Frictions? Parsing the Evidence from the Financial Crisis of 2007 to 2009," *NBER Macroeconomics Annual*, 2013, *27* (1), 159 – 214.

Agarwal, Sumit and Artashes Karapetyan, "Salience and Mispricing: Homebuyers' Housing Decisions," *Norges Bank Working Paper 21*, 2016.

Allcott, Hunt, "Consumers' Perceptions and Misperceptions of Energy Costs," *American Economic Review*, May 2011, *101* (3), 98–104.

Allen, Franklin and Douglas Gale, "Limited Market Participation and Volatility of Asset Prices," *The American Economic Review*, 1994, pp. 933–955.

Allen, Linda and Aron A. Gottesman, "The Informational Efficiency of the Equity Market as Compared to the Syndicated Bank Loan Market," *Journal of Financial Services Research*, 2006, *30* (1), 5–42.

Almeida, Heitor, Murillo Campello, Bruno Laranjeira, and Scott Weisbenner, "Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis," *Critical Finance Review*, 2011, *1*, 3–58.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection," *Journal of the European Economic Association*, 2003, *1* (1), 68–123.

Andonov, Aleksandar and Joshua D. Rauh, "The Return Expectations of Institutional Investors," *Working Paper*, 2020.

Anton, Miguel and Christopher Polk, "Connected Stocks," *The Journal of Finance*, 2014, *69* (3), 1099–1127.

Arrow, Kenneth J., "The Theory of Risk Aversion," in "Aspects of The Theory of Risk Bearing" 1965.

Asch, Solomon E., "Forming Impressions of Personality.," *The Journal of Abnormal and Social Psychology*, 1946, *41* (3), 258.

Back, Kerry E., *Asset Pricing and Portfolio Choice Theory*, Oxford University Press, 2017.

Baker, Malcolm and Jeffrey Wurgler, "The Equity Share in New Issues and Aggregate Stock Returns," *The Journal of Finance*, 2000, *55* (5), 2219–2257.

_, Robin Greenwood, and Jeffrey Wurgler, "The Maturity of Debt Issues and Predictable Variation in Bond Returns," *Journal of Financial Economics*, 2003, *70* (2), 261–291.

Bao, Jack, Jun Pan, and Jiang Wang, "The Illiquidity of Corporate Bonds," *The Journal of Finance*, 2011, *66* (3), 911–946.

Barber, Brad M., Xing Huang, and Terrance Odean, "Which Factors Matter to Investors? Evidence from Mutual Fund Flows," *The Review of Financial Studies*, 2016, *29* (10), 2600–2642.

Barclay, Michael J., Neil D. Pearson, and Michael S. Weisbach, "Open-end Mutual Funds and Capital-gains Taxes," *Journal of Financial Economics*, 1998, *49* (1), 3 – 43.

Becker, Bo and Victoria Ivashina, "Cyclicality of Credit Supply: Firm Level Evidence," *Journal of Monetary Economics*, 2014, *62*, 76–93.

Benmelech, Efraim, Jennifer Dlugosz, and Victoria Ivashina, "Securitization Without Adverse Selection: The case of CLOs," *Journal of Financial Economics*, 2012, *106* (1), 91–113.

Benos, Evangelos, Marek Jochec, and Victor Nyekel, "Can Mutual Funds Time Risk Factors?," *The Quarterly Review of Economics and Finance*, 2010, *50* (4), 509 – 514.

Bergstresser, Daniel and James Poterba, "Do After-tax Returns Affect Mutual Fund Inflows?," *Journal of Financial Economics*, 2002, *63* (3), 381 – 414.

Berk, Jonathan B. and Jules H. van Binsbergen, "Assessing Asset Pricing Models Using Revealed Preference," *Journal of Financial Economics*, 2016, *119* (1), 1–23.

Besley, Timothy, Neil Meads, and Paolo Surico, "The Incidence of Transaction Taxes: Evidence from a Stamp Duty Holiday," *Journal of Public Economics*, 2014, *119* (C), 61–70.

Best, Michael and Henrik Jacobsen Kleven, "Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the U.K," *Review of Economic Studies*, 2018, *85* (1), 157–193.

Beyhaghi, Mehdi and Sina Ehsani, "The Cross-Section of Expected Returns in the Secondary Corporate Loan Market," *The Review of Asset Pricing Studies*, 2017, *7* (2), 243–277.

Black, Fisher and Robert Litterman, "Global Portfolio Optimization," *Financial Analysts Journal*, 1992, *48* (05), 28–43.

Bozanic, Zahn, Maria Loumioti, and Florin P. Vasvari, "Corporate Loan Securitization and the Standardization of Financial Covenants," *Journal of Accounting Research*, 2018, *56* (1), 45–83.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs," *R Journal*, 2015, *7* (1), 38.

Camerer, Colin and Teck Hua Ho, "Experience-weighted Attraction Learning in Normal Form Games," *Econometrica*, 1999, *67* (4), 827–874.

Cattaneo, Matia D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimator," *Journal of the American Statistical Association*, 2020.

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng, "On Binscatter," *arXiv preprint arXiv:1902.09608*, 2019.

Chernenko, Sergey and Adi Sunderam, "Liquidity Transformation in Asset Management: Evidence From the Cash Holdings of Mutual Funds," Technical Report, National Bureau of Economic Research 2016.

_ , Samuel G. Hanson, and Adi Sunderam, "Who Neglects Risk? Investor Experience and the Credit Boom," *Journal of Financial Economics*, 2016, *122* (2), 248–269.

Chetty, Raj, "A New Method of Estimating Risk Aversion," *American Economic Review*, 2006, *96* (5), 1821–1834.

_ , Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, September 2009, *99* (4), 1145–1177.

Chiang, Yao-Min, David A. Hirshleifer, Yiming Qian, and Ann E. Sherman, "Do Investors Learn from Experience? Evidence from Frequent IPO Investors," *The Review of Financial Studies*, 2011, *24* (5), 1560–1589.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, "Reinforcement Learning and Savings Behavior," *The Journal of Finance*, 2009, *64* (6), 2515–2534.

Cochrane, John H., "Macro-Finance," *Review of Finance*, 2017, *21* (3), 945–985.

Cohen, Gregory, Melanie Friedrichs, Kamran Gupta, William Hayes, Seung Jung Lee, W Blake Marsh, Nathan Mislang, Maya Shaton, and Martin Sicilian, "The US Syndicated Loan Market: Matching Data," *Federal Reserve Bank of Kansas City Working Paper*, 2018, (18-09).

Cohen, Randolph B., Christopher Polk, and Bernhard Silli, "Best Ideas," LSE Research Online Documents on Economics 24471, London School of Economics and Political Science, LSE Library 2008.

Colla, Paolo, Filippo Ippolito, and Kai Li, "Debt Specialization," *The Journal of Finance*, 2013, *68* (5), 2117–2141.

Cordell, Larry, Michael R. Roberts, and Michael Schwert, "CLO Performance," 2020.

Coval, Joshua and Erik Stafford, "Asset Fire Sales (and Purchases) in Equity Markets," *Journal of Financial Economics*, 2007, *86* (2), 479–512.

CreditFlux, "Guide to CLOs," 2015.

Cvitanic, Jaksa and Ioannis Karatzas, "Convex Duality in Constrained Portfolio Optimization," *The Annals of Applied Probability*, 1992, *2* (4), 767–818.

Davidson, Russell and James G. MacKinnon, "Artificial Regressions," *A Companion to Theoretical Econometrics*, 2001, pp. 16–37.

Deese, James and Roger A. Kaufman, "Serial Effects in Recall of Unorganized and Sequentially Organized Verbal Material.," *Journal of Experimental Psychology*, 1957, *54* (3), 180.

Dong, Ming, David Hirshleifer, and Siew Hong Teoh, "Overvalued Equity and Financing Decisions," *The Review of Financial Studies*, 2012, *25* (12), 3645–3683.

Duffie, Darrell, "Presidential Address: Asset Price Dynamics with Slow-Moving Capital," *The Journal of Finance*, 2010, *65* (4), 1237–1267.

Ebbinghaus, Hermann, "Memory: A Contribution to Experimental Psychology," 1913, p. 142.

Elkamhi, Redouane and Yoshio Nozawa, "Too Big to Diversify: A Stress Test on Collateralized Loan Obligations," *Available at SSRN 3635086*, 2021.

Ellul, Andrew, Chotibhak Jotikasthira, Christian T. Lundblad, and Yihui Wang, "Mark-to-Market Accounting and Systemic Risk: Evidence from the Insurance Industry," *Economic Policy*, 11 2014, *29* (78), 297–341.

_ , _ , _ , and _ , "Is Historical Cost Accounting a Panacea? Market Stress, Incentive Distortions, and Gains Trading," *The Journal of Finance*, 2015, *70* (6), 2489–2538.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, "Explaining the Rate Spread on Corporate Bonds," *The Journal of Finance*, 2001, *56* (1), 247–277.

Erev, Ido and Alvin E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 1998, pp. 848–881.

European Commission, "Taxation Trends in the European Union," *DG Taxation and Customs Union*, 2018.

Evans, Richard B., "Mutual Fund Incubation," *The Journal of Finance*, 2010, *65* (4), 1581–1611.

Fabozzi, Frank, Sven Klingler, Pia Mølgaard, and Mads Stenbo Nielsen, "Active Loan Trading," *Journal of Financial Intermediation*, 2020, p. 100868.

Fama, Eugene F. and Kenneth R. French, "A Five-factor Asset Pricing Model," *Journal of Financial Economics*, 2015, *116* (1), 1–22.

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *American Economic Review*, 2020, *110* (1), 298–336.

Faulkender, Michael and Mitchell A. Petersen, "Does the Source of Capital Affect Capital Structure?," *The Review of Financial Studies*, 2006, *19* (1), 45–79.

Fazio, Russell, Mark Zanna, and Joel Cooper, "Direct Experience and Attitude-Behavior Consistency: An Information Processing Analysis," *Personality and Social Psychology Bulletin*, 1978, *4*, 48–51.

Fei, Fei, Ana-Maria Fuertes, and Elena Kalotychou, "Credit Rating Migration Risk and Business Cycles," *Journal of Business Finance & Accounting*, 2012, *39* (1-2), 229–263.

Feldman, Naomi E. and Bradley J. Ruffle, "The Impact of Including, Adding, and Subtracting a Tax on Demand," *American Economic Journal: Economic Policy*, February 2015, *7* (1), 95–118.

Feldman, Naomi, Jacob Goldin, and Tatiana Homonoff, "Raising the Stakes: Experimental Evidence on the Endogeneity of Taxpayer Mistakes," *National Tax Journal*, June 2018, *71* (2), 201–230.

Financial Stability Board, "Vulnerabilities Associated with Leveraged Loans and Collateralised Loan Obligations," 2019.

Finkelstein, Amy, "E-ztax: Tax Salience and Tax Rates," *The Quarterly Journal of Economics*, 2009, *124* (3), 969–1010.

Fleckenstein, Quirin, Manasa Gopal, German Gutierrez Gallardo, and Sebastian Hillenbrand, "Nonbank Lending and Credit Cyclicality," *Available at SSRN 3629232*, 2020.

Frank, Murray Z. and Vidhan K. Goyal, "Capital Structure Decisions: which Factors are Reliably Important?," *Financial Management*, 2009, *38* (1), 1–37.

Frazzini, Andrea and Owen A. Lamont, "Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns," *Journal of Financial Economics*, 2008, *88* (2), 299–322.

Gande, Amar and Anthony Saunders, "Are Banks Still Special When There is a Secondary Market for Loans?," *The Journal of Finance*, 2012, *67* (5), 1649–1684.

Gao, Pengjie and Dong Lou, "Cross-Market Timing in Security Issuance," in "AFA 2012 Chicago Meetings Paper" 2013.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five Facts About Beliefs and Portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

Goldfarb, Sam, "Wall Street's Answer to Risks in Loan Market: Bundle Lower-Rated Loans," *Wall Street Journal*, 2019.

Goldin, Jacob, "Optimal Tax Salience," *Journal of Public Economics*, 2015, *131* (C), 115–123.

Goldstein, Itay, Hao Jiang, and David T. Ng, "Investor Flows and Fragility in Corporate Bond Funds," *Journal of Financial Economics*, 2017, *126* (3), 592–613.

Greenwood, Robin and Andrei Shleifer, "Expectations of Returns and Expected Returns," *The Review of Financial Studies*, 2014, *27* (3), 714–746.

_ and Samuel G. Hanson, "Issuer Quality and Corporate Bond Returns," *The Review of Financial Studies*, 2013, *26* (6), 1483–1525.

_ and Stefan Nagel, "Inexperienced Investors and Bubbles," *Journal of Financial Economics*, 2009, *93* (2), 239–258.

Harford, Jarrad, Marc Martos-Vila, and Matthew Rhodes-Kropf, *Corporate Financial Policies in Misvalued Credit Markets*, Citeseer, 2014.

Harrison, J. Michael and David M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 1979, *20* (3), 381–408.

Hirshleifer, David A., Ben Lourie, Thomas Ruchti, and Phong Truong, "First Impressions and Analyst Forecast Bias," *Review of Finance*, 2021, *25* (2), 325–364.

Hoffmann, Arvid O. I., Zwetelina Iliewa, and Lena Jaroszek, "Wall Street Crosses Memory Lane: How Witnessed Returns Affect Professionals' Expected Returns," in "Paris December 2017 Finance Meeting EUROFIDAI-AFFI" 2017.

Ivashina, Victoria and Zheng Sun, "Institutional Demand Pressure and the Cost of Corporate Loans," *Journal of Financial Economics*, 2011, *99* (3), 500–522.

_ and _ , "Institutional Stock Trading on Loan Market Information," *Journal of financial Economics*, 2011, *100* (2), 284–303.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, "Unobserved Actions of Mutual Funds," *The Review of Financial Studies*, 2006, *21* (6), 2379–2416.

Kaustia, Markku and Samuli Knüpfer, "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions," *The Journal of Finance*, 2008, *63* (6), 2679–2702.

Knüpfer, Samuli, Elias Rantapuska, and Matti Sarvimäki, "Formative Experiences and Portfolio Choice: Evidence from the Finnish Great Depression," *The Journal of Finance*, 2017, *72* (1), 133–166.

Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 1996, *34* (1), 42–71.

Koijen, Ralph S.J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Kothari, S.P., Dalia Blass, Alan Cohen, and Sumit Rajpal, "U.S. Credit Markets: Interconnectedness and the Effects of the COVID-19 Economic Shock," *U.S. Security and Exchange Commission Report*, 2020.

Kotlikoff, Laurence and Lawrence Summers, "Tax Incidence," in A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, 1 ed., Vol. 2, Elsevier, 1987, chapter 16, pp. 1043–1092.

Kundu, Shoini, "The Externalities of Fire Sales: Evidence from Collateralized Loan Obligations," *Working Paper*, 2021.

Leary, Mark T., "Bank Loan Supply, Lender Choice, and Corporate Capital Structure," *The Journal of Finance*, 2009, *64* (3), 1143–1185.

Ledoit, Olivier and Michael Wolf, "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, 2004, *30*, 110–119.

Lemmon, Michael and Michael R. Roberts, "The Response of Corporate Financing and Investment to Changes in the Supply of Credit," *Journal of Financial and Quantitative Analysis*, 2010, pp. 555–587.

Liebscher, Roberto and Thomas Mählmann, "Are Professional Investment Managers Skilled? Evidence from Syndicated Loan Portfolios," *Management Science*, 2017, *63* (6), 1892–1918.

Lochstoer, Lars A. and Tyler Muir, "Volatility Expectations and Returns," *The Journal of Finance*, forthcoming.

Lou, Dong, "A Flow-based Explanation for Return Predictability," *The Review of Financial Studies*, 2012, *25* (12), 3457–3489.

Loughran, Tim and Jay R. Ritter, "The New Issues Puzzle," *The Journal of Finance*, 1995, *50* (1), 23–51.

Loumioti, Maria and Florin P. Vasvari, "Consequences of CLO Portfolio Constraints," *Available at SSRN 3371162*, 2019.

_ and _ , "Portfolio Performance Manipulation in Collateralized Loan Obligations," *Journal of Accounting and Economics*, 2019, *67* (2-3), 438–462.

Ma, Yueran, "Nonfinancial Firms as Cross-Market Arbitrageurs," *The Journal of Finance*, 2019, *74* (6), 3041–3087.

Maenhout, Pascal J., "Robust Portfolio Rules and Asset Pricing," *The Review of Financial Studies*, 2004, *17* (4), 951–983.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

_ and _ , "Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

_ , _ , and Zhen Yan, "The Making of Hawks and Doves," *Journal of Monetary Economics*, 2021, *117*, 19–42.

Martin, Ian W. R., "What is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

Mehra, Rajnish and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Merton, Robert C., "On Estimating the Expected Return on the Market : An Exploratory Investigation," *Journal of Financial Economics*, 1980, *8* (4), 323–361.

Moreira, Alan and Tyler Muir, "Should Long-Term Investors Time Volatility?," *Journal of Financial Economics*, 2019, *131* (3), 507 – 527.

Morningstar, "CLO Commentary: Push for Higher Equity Returns Leads to Weaker Structural Features in CLOs," 2018.

Morris, Stephen, Ilhyock Shim, and Hyun Song Shin, "Redemption Risk and Cash Hoarding by Asset Managers," *Journal of Monetary Economics*, 2017, *89*, 71–87.

Murdock, Bennet B. Jr., "The Serial Position Effect of Free Recall.," *Journal of Experimental Psychology*, 1962, *64* (5), 482.

_ , *Human Memory: Theory and Data.*, Lawrence Erlbaum, 1974.

Murfin, Justin R. and Ryan Pratt, "Comparables Pricing," *The Review of Financial Studies*, 04 2018, *32* (2), 688–737.

Nadauld, Taylor D. and Michael S. Weisbach, "Did Securitization Affect the Cost of Corporate Debt?," *Journal of Financial Economics*, 2012, *105* (2), 332–352.

Neyman, Jerzy and Elizabeth L. Scott, "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 1948, pp. 1–32.

Ngai, L Rachel and Silvana Tenreyro, "Hot and cold seasons in the housing market," *American Economic Review*, 2014, *104* (12), 3991–4026.

Nipher, Francis E., "On the Distribution of Errors in Numbers Written from Memory," *Transactions of the Academy of Science of St. Louis*, 1878, *3*, 10–1.

Nofsinger, John R. and Richard W. Sias, "Herding and Feedback Trading by Institutional and Individual Investors," *The Journal of Finance*, 1999, *54* (6), 2263–2295.

Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica*, 1964, *32*, 122–136.

Reck, Daniel, "Taxes and Mistakes: What's in a Sufficient Statistic?," *Working Paper*, April 2016.

Roll, Richard, "A Simple Implicit Measure of the Effective Bid-ask Spread in an Efficient Market," *The Journal of Finance*, 1984, *39* (4), 1127–1139.

Roth, Alvin E. and Ido Erev, "Learning in Extensive-form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 1995, *8* (1), 164–212.

Schmidt, Lawrence, Allan Timmermann, and Russ Wermers, "Runs on Money Market Mutual Funds," *American Economic Review*, 2016, *106* (9), 2625–57.

Sharpe, William F., "Imputing Expected Security Returns from Portfolio Composition," *Journal of Financial and Quantitative Analysis*, 1974, *9* (03), 463–472.

Shleifer, Andrei and Robert Vishny, "Fire Sales in Finance and Macroeconomics," *Journal of Economic Perspectives*, 2011, *25* (1), 29–48.

Shumway, Tyler, Maciej J. Szefler, and Kathy Yuan, "The Information Content of Revealed Beliefs in Portfolio Holdings," LSE Research Online Documents on Economics, London School of Economics and Political Science, LSE Library 2011.

Sialm, Clemens and Hanjiang Zhang, "Tax-Efficient Asset Management: Evidence from Equity Mutual Funds," *The Journal of Finance*, 2020, *75* (2), 735–777.

Simonsohn, Uri, Niklas Karlsson, George Loewenstein, and Dan Ariely, "The Tree of Experience in the Forest of Information: Overweighing Experienced Relative to Observed Information," *Games and Economic Behavior*, 2008, *62* (1), 263–286.

S&P Global, "As Leveraged Loan Downgrades Mount, CLOs Cast Wary Eye on Triple-C Limits," 2019.

__ , "Default, Transition, and Recovery: 2019 Annual Global Corporate Default And Rating Transition Study," 2020.

__ , "Leveraged Commentary & Data (LCD): Leveraged Loan Primer," 2020.

Standard & Poor's Compustat Services, I., "Compustat," 1962.

Sufi, Amir, "The Real Effects of Debt Certification: Evidence from the Introduction of Bank Loan Ratings," *The Review of Financial Studies*, 2009, *22* (4), 1659–1691.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Tepla, Lucie, "Optimal Portfolio Policies with Borrowing and Shortsale Constraints," *Journal of Economic Dynamics and Control*, 2000, *24* (11-12), 1623–1639.

Thomson Financial, I., "SDC Platinum," 2001.

Thorndike, Edward L., "Animal Intelligence: An Experimental Study of the Associative Processes in Animals.," *The Psychological Review: Monograph Supplements*, 1898, *2* (4), i.

Touloumis, Anestis, "Nonparametric Stein-type Shrinkage Covariance Matrix Estimators in High-dimensional Settings," *Computational Statistics & Data Analysis*, 2015, *83* (C), 251–261.

Tversky, Amos and Daniel Kahneman, "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 1973, *5* (2), 207–232.

University of Chicago. Center for Research in Security Prices, I., "CRSP databases," 1960.

van Binsbergen, Jules H., Michael W. Brandt, and Ralph S.J. Koijen, "Optimal Decentralized Investment Management," *The Journal of Finance*, 2008, *63* (4), 1849–1895.

Warther, Vincent A., "Aggregate Mutual Fund Flows and Security Returns," *Journal of Financial Economics*, 1995, *39* (2-3), 209–235.

Wermers, Russ, "Mutual Fund Herding and the Impact on Stock Prices," *The Journal of Finance*, 1999, *54* (2), 581–622.

Xu, Gan-Lin and Steven E. Shreve, "A Duality Method for Optimal Consumption and Investment Under Short- Selling Prohibition. I. General Market Coefficients," *The Annals of Applied Probability*, 1992, *2* (1), 87–112.

Zhu, Qifei, "Capital supply and Corporate Bond Issuances: Evidence from Mutual Fund Flows," *Journal of Financial Economics*, forthcoming.

A. Appendix to Contagion in the Market for Leveraged Loans

A.1 Tables

Table A1: Summary Statistics - Holdings

The table contains summary statistics for the sample of CLOs in the CLO-i dataset between January 2009 and December 2019. Panel A reports summary statistics at the level of each CLO report, while Panel B aggregates at the level of Management Team - Month. Total Assets refers to the sum of the current balance of securities held measured in $Mlns; Nr. Issuers refers to the distinct number of issuers held; Nr. Securities to the distinct number of securities; % of Assets to the fraction of assets represented by each security in the portfolio; Interest rate to the interest rate of the loan; % CCC to the ratio between the sum of the current balance of securities rated at or below Caa (CCC) and total assets; % Default to the ratio between the sum of the current balance of securities in default and total assets; Age to the difference between the first time a certain deal or management team appears in sample and the current reporting date; WARF to the weighted-average rating factor computed using Moody's rating factors.

	Nr.Obs	Min	Max	Median	Mean	Std.Dev
Panel A: CLO Deals						
Total Assets	89,111	0.00	18336.84	408.86	434.73	283.96
Nr. Issuers	89,111	1.00	748.00	157.00	168.20	101.33
Nr. Securities	89,111	1.00	907.00	203.00	213.82	115.27
% of Assets	88,517	0.00	1.00	0.01	0.01	0.05
Interest Rate	88,334	0.00	16.47	4.77	4.87	1.11
% CCC	88,905	0.00	1.00	0.06	0.09	0.12
% Default	88,905	0.00	1.00	0.02	0.05	0.11
Age	89,111	0.00	10.90	2.00	2.57	2.16
WARF	88,421	4.33	10000.00	2737	2923.48	806.63
Panel B: Management Teams	Nr.Obs	Min	Max	Median	Mean	Std.Dev
Deals Managed	15,470	1.00	53.00	3.00	5.57	6.50
Total Assets	15,470	0.00	28504.00	1239.46	2504.18	3328.16
Nr. Issuers	15,470	1.00	1209.00	217.00	257.45	183.73
Nr. Securities	15,470	1.00	4935.00	337.00	565.32	606.29
% of Assets	15,100	0.00	1.00	0.00	0.01	0.04
Interest Rate	15,298	0.00	12.28	4.72	4.88	1.10
% CCC	15,451	0.00	1.00	0.07	0.09	0.10
% Default	15,451	0.00	1.00	0.03	0.05	0.09
Age	15,470	0.00	11.00	4.21	4.70	3.23
WARF	15,080	160.75	10000.00	2753.89	2944.44	818.51

Table A2: Summary Statistics - Transactions

The table contains summary statistics for the sample of CLOs' transactions in the CLO-i dataset between January 2009 and December 2019. Tot. Transactions refers to the number of transactions completed by a deal between two reporting dates; Purchase Price to the average price at which a loan has been purchased; Nr. Purchases to the total number of securities bought; Amt. Purchased to the amount of securities purchased, measured in $ Mlns; Sale Price to the average price at which a loan has been sold; Nr. Sales to the total number of securities sold; Amt. Sold to the amount of securities sold, measured in $ Mlns. Prices have been capped between $10 and $150. Nr. Obs counts the number of non-zero observations, while Nr. Zeros counts the number of observations equal to zero. Min, Max, Median, Mean and Std. Dev are the minimum, maximum, median, average and standard deviation of the non-zero observations.

	CLO Deals						
	Nr. Obs	Nr. Zeros	Min	Max	Median	Mean	Std.Dev
Tot. Transactions	72,515	16,596	0.00	853.00	19.00	29.73	36.97
Purchase Price	64,919	0	10.00	150.00	98.99	95.31	13.39
Nr. Purchases	65,458	23,653	0.00	736.00	13.00	19.25	26.20
Amt. Purchased	65,458	23,653	0.00	2372.24	15.29	22.44	43.94
Sale Price	64,755	0	10.00	150.00	98.20	91.73	17.53
Nr. Sales	65,239	23,872	0.00	475.00	8.00	13.69	17.55
Amt. Sold	65,239	23,872	0.00	1006.72	6.65	10.84	19.67

Table A3: The Mechanical Effect of Downgrades to Caa (CCC) on OC Tests

The table studies the mechanical effect of downgrades to Caa (CCC) on the slack of OC tests. Columns (1)-(4) report the results of the following regression: $\text{slack}^k_{i,t} = \alpha + \beta_1 \text{Shocked}_{i,t} + X_{i,t}\delta + \varepsilon_{i,t}$, where $\text{slack}^k_{i,t} = \frac{\widetilde{OC}^k_{i,t} - O^k_i}{O^k_i}$, $\widetilde{OC}^k_{i,t}$ is the realization of the OC test for tranche k and CLO i and O^k_i is the test threshold; $\text{Shocked}_{i,t}$ is an indicator variable that turns on whenever the loans of CLO i have been downgraded to Caa (CCC). Column (5) reports the result of the following regression: $\text{slack}^k_{i,t} = \alpha + \beta_1 \text{Shocked}_{i,t} + \beta_2 \text{Above 7.5\%}_{i,t} + \beta_3 \text{Shocked}_{i,t} \times \text{Above 7.5\%}_{i,t} + X_{i,t}\delta + \varepsilon_{i,t}$, where $\text{Above 7.5\%}_{i,t}$ is a dummy variable that turns on whenever the fraction of Caa (CCC) securities for CLO i is greater than 7.5% at time t. Standard errors are reported in parantheses and are double clustered at the Year × Month & CLO Deal level.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(Intercept)	0.072***				
	(0.008)				
Shocked	-0.007	-0.007	-0.008	-0.008	-0.006
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
Above 7.5%					0.140
					(0.135)
Shocked × Above 7.5%					-0.124
					(0.128)
Fixed-Effects					
Year×Month	No	Yes	Yes	Yes	Yes
Deal	No	No	Yes	Yes	Yes
Senior/Junior OC	No	No	No	Yes	Yes
Observations	80,321	80,321	80,321	80,321	80,321
R^2	0.000	0.001	0.052	0.053	0.053
Within R^2	–	0.000	0.000	0.000	0.000

Two-way (Year×Month & Deal) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A4: Holdings vs. Market Prices

The table compares the discount of loans as they are reported by CLOs with the closest market price in the following year by running the following regression: $\text{discount}_{j,t} = \alpha_i + \alpha_t + \beta_1 \text{Transaction}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$ is the discount at which a loan is recorded or traded compared to par, $\text{Transaction}_{j,t}$ is an indicator variable equal to one whenever the price comes from an actual sale transaction and zero otherwise, α_i and α_t are issuer and time fixed effects, while $X_{j,t}$ includes a set of fixed effects for rating, industry and interest rate of the loan. Panel A consider the universe of loans that have at least one price reported on the balance sheet of CLOs and one market price in the following twelve months, while Panel B focuses on the subset of loans rated Caa (CCC). Standard errors clustered at the Year ×Month and Issuer level are reported in parentheses.

Panel A: All Loans	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Transaction	0.871***	0.875***	0.875***	0.875***	0.878***
	(0.169)	(0.170)	(0.170)	(0.170)	(0.170)
Fit statistics					
Obs.	2,211,675	2,200,337	2,200,337	2,200,337	2,211,675
R^2	0.578	0.609	0.610	0.610	0.726
Within R^2	0.004	0.005	0.005	0.005	0.006
Panel B: Caa (CCC) Loans	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Transaction	3.990***	4.000***	4.000***	4.000***	4.030***
	(0.801)	(0.801)	(0.801)	(0.801)	(0.803)
Fit statistics					
Obs.	195,073	195,073	195,073	195,073	195,073
R^2	0.622	0.644	0.646	0.646	0.728
Within R^2	0.013	0.014	0.014	0.014	0.018
Fixed-Effects					
Year × Month	Yes	Yes	Yes	Yes	No
Issuer	Yes	Yes	Yes	Yes	No
Rating	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Interest	No	No	No	Yes	Yes
Year × Month×Issuer	No	No	No	No	Yes

Two-way (Year× Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5: Discontinuity Test

The table reports the result of a discontinuity test where I compare the density of the fraction of Caa (CCC) or lower rated securities before and after the 7.5% threshold. Number of Obs. reports the total number of observations on the left and right of the 7.5% threshold; Eff. Number of Obs. reports the number of observations used for the test which employs Cattaneo et al. (2020) local polynomial approximation.

Cutoff 0.075	Left	Right
Number of Obs.	45318	1620
Eff. Number of Obs.	961	744
Order	2	2
Order Bias	3	3
Bandwith	0.008	0.011
	T	Pr $> \|T\|$
Statistic	-0.2725	0.7852

Table A6: Par Building

Columns (1) and (2) report the results of the following regressions: $\text{gain}_{i,j,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,j,t} = 100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for purchases and $\text{gain}_{i,j,t} = -100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for sales; $\text{Constrained}_{i,t}$ is a dummy variable equal to one whenever the Junior (column (1)) or Senior (column(2)) slack of CLO i is between 0% and and 5% in period t; $\text{Shocked}_{i,t}$ is a dummy variable equal to one whenever the loans of CLO i have been downgraded. Columns (3) and (4) report the results of the following regressions: $\text{gain}_{i,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,t} = 100 \times \left(\sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ and the other variables are defined as above. $\text{Constrained}_{i,t}$ refers to Junior tests in column (3) and to Senior tests in column (4). Standard errors are reported in parentheses and they are double clustered at the Year×Month & CLO Deal level.

	Individual Transactions		**Multiple Transactions**	
	(1)	**(2)**	**(3)**	**(4)**
(Intercept)	-0.007***	-0.003***	-0.068***	-0.034***
	(0.001)	(0.000)	(0.006)	(0.002)
Constrained	0.004***	-0.008***	0.035***	-0.052***
	(0.001)	(0.002)	(0.006)	(0.015)
Shocked	0.006***	0.005***	0.050***	0.033***
	(0.001)	(0.000)	(0.006)	(0.003)
Constrained × Shocked	0.005***	0.004	0.069***	0.016
	(0.001)	(0.003)	(0.008)	(0.022)
Fit statistics				
Observations	309,028	303,160	30,156	29,034
R^2	0.002	0.002	0.009	0.005
Adjusted R^2	0.002	0.002	0.009	0.005
OC Test	Junior	Senior	Junior	Senior

Two-way (Year×Month & CLO Deal) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A7: Rating Factor

The table shows how to convert Moody's and Standard&Poor's ratings into Moody's Rating Factors. Rating Factors convert the ordinal rating into a cardinal variable.

Moody's	S&P	Rating Factor	Moody's	S&P	Rating Factor
Aaa	AAA	1	Ba1	BB+	940
Aa1	AA+	10	Ba2	BB	1350
Aa2	AA	20	Ba3	BB-	1766
Aa3	AA-	40	B1	B+	2220
A1	A+	70	B2	B	2720
A2	A	120	B3	B-	3490
A3	A-	180	Caa1	CCC+	4770
Baa1	BBB+	260	Caa2	CCC	6500
Baa2	BBB	360	Caa3	CCC-	8070
Baa3	BBB-	610	Ca-C	CC-C	10000

Table A8: WARF Deterioration

The table compares the average rating factor for loans sold, in column (1), and purchased, column (2), by CLOs whose loans have been downgraded to Caa (CCC) and whose Junior OC tests are binding, by reporting the coefficients of the following regression: $\text{RF}_{i,j,t} = \alpha + \beta_1 \text{Shocked}_{i,t} + \beta_2 \text{Constrained}_{i,t} + \beta_3 \text{Shocked}_{i,t} \times \text{Constrained}_{i,t} + \varepsilon_{i,t}$; where $\text{RF}_{i,j,t}$ is the rating factor of loan j, sold by CLO i at time t; $\text{Shocked}_{i,t}$ is dummy variable equal to one when the loans of CLO i have been downgraded to Caa (CCC); $\text{Constrained}_{i,t}$ is a dummy variable equal to one when the slack of the Junior OC test is between 0% and 5%. Column (3) aggregates the results by reporting the coefficients of the following regression: $\Delta\text{WARF}_{i,t} = \alpha + \beta_1 \text{Shocked}_{i,t} + \beta_2 \text{Constrained}_{i,t} + \beta_3 \text{Shocked}_{i,t} \times \text{Constrained}_{i,t} + \varepsilon_{i,t}$, where $\Delta\text{WARF}_{i,t} = \sum_j \text{RF}_{i,j,t} \times \frac{\text{Amt. Purchased}_{i,j,t}}{\sum_j \text{Amt. Purchased}_{i,j,t}} - \sum_j \text{RF}_{i,j,t} \times \frac{\text{Amt. Sold}_{i,j,t}}{\sum_j \text{Amt. Sold}_{i,j,t}}$. Standard errors clustered by CLO Deal & Year×Month are reported in parentheses.

	Individual Transactions		**Multiple Transactions**
	(1)	**(2)**	**(3)**
(Intercept)	3025.0***	2641.4***	-441.3***
	(29.5)	(12.3)	(39.6)
Shocked	-151.6***	67.5***	316.2***
	(26.9)	(18.3)	(35.5)
Constrained	73.9**	-6.9	374.5***
	(34.2)	(14)	(43.3)
Shocked × Constrained	-514.2***	105.8***	79.1*
	(40.2)	(23.4)	(41.2)
Fit statistics			
Observations	155,079	162,629	21,043
R^2	0.00847	0.00434	0.04243
Adjusted R^2	0.00845	0.00432	0.0423

Two-way (CLO Deal & Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A9: Price Pressure

The table reports the results of the following regression: $\text{discount}_{j,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,k,t} = 100 \times \log(100/P_{j,k,t})$, $P_{j,k,t}$ is the price of loan j issued by firm k at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by shocked CLOs is above median, $\text{Post}_{j,t}$ is a dummy equal to one after loan j has received an above median selling volume by shocked CLOs, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. All the regressions include the lagged average discount on the issuer computed as $\text{Avg. discount}_{k,t-1} = \frac{1}{J_k \times (t-1)} \sum_{j=1}^{J_k} \sum_{s=1}^{t-1} \text{discount}_{j,k,s}$, where J_k is the number of loans by issuer k actively traded. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked×Post	0.718***	0.494***	0.476***	0.469***	0.436***
	(0.130)	(0.104)	(0.088)	(0.083)	(0.086)
Shocked	-0.395***	-0.093	-0.009	-0.027	-0.011
	(0.116)	(0.094)	(0.068)	(0.063)	(0.067)
Avg. discount_{t-1}	0.802***	0.798***	0.659***	0.651***	0.660***
	(0.037)	(0.037)	(0.033)	(0.032)	(0.033)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	738,354	738,354	738,354	738,354	738,354
R^2	0.421	0.432	0.533	0.564	0.540
Within R^2	0.324	0.312	0.224	0.218	0.223

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A10: Price Pressure Within Issuers

The table reports the results of the following regression: $\text{discount}_{j,k,t} = \alpha_{k,t} + \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,k,t}$, where $\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$, $P_{j,t}$ is the price of loan j at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by shocked CLOs is above median, $\text{Post}_{j,t}$ is a dummy equal to one after loan j has received an above median selling volume by shocked CLOs, $X_{j,t}$ is a matrix containing the fixed effects reported in the table. Column (1) includes all the loans that have been traded by CLOs; column (2) restricts the sample to those issuers with at least two actively traded loans; column (3) to those issuers with at least five actively traded loans; column (4) to those issuers with at least 10 actively traded loans. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	(2)	(3)	(4)
Shocked	-0.032	-0.034	-0.053*	-0.042
	(0.026)	(0.026)	(0.032)	(0.040)
Shocked×Post	0.120***	0.126***	0.153***	0.179**
	(0.039)	(0.041)	(0.051)	(0.071)
Fixed-effects				
Year×Month×Issuer	Yes	Yes	Yes	Yes
Year×Month×TTM	Yes	Yes	Yes	Yes
Year×Month×Rating	Yes	Yes	Yes	Yes
Year×Month×Industry	Yes	Yes	Yes	Yes
Year×Month×Interest	Yes	Yes	Yes	Yes
Fit statistics				
Observations	746,956	629,507	402,252	170,122
R^2	0.914	0.884	0.853	0.821
Within R^2	0.000	0.000	0.000	0.000
Nr. Traded Loans	≥ 1	≥ 2	≥ 5	≥ 10

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A11: Price Pressure - Purchases

The table reports the results of the following regression: $\text{discount}_{j,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,k,t} = 100 \times \log(100/P_{j,k,t})$, $P_{j,k,t}$ is the price of loan j issued by firm k at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j purchasing volume by shocked CLOs is above median, $\text{Post}_{j,t}$ is a dummy equal to one after loan j has received an above median purchasing volume by shocked CLOs, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. All the regressions include the lagged average discount on the issuer computed as $\text{Avg. discount}_{k,t-1} = \frac{1}{J_k \times (t-1)} \sum_{j=1}^{J_k} \sum_{s=1}^{t-1} \text{discount}_{j,k,s}$, where J_k is the number of loans by issuer k actively traded. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-1.16***	-0.923***	-0.612***	-0.634***	-0.603***
	(0.130)	(0.109)	(0.084)	(0.086)	(0.084)
Shocked×Post	-0.108	-0.281**	0.014	0.004	-0.054
	(0.149)	(0.142)	(0.139)	(0.135)	(0.133)
Avg. discount_{t-1}	0.795***	0.790***	0.611***	0.609***	0.607***
	(0.036)	(0.037)	(0.039)	(0.038)	(0.038)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	738,354	738,354	597,976	597,976	596,807
R^2	0.42337	0.42419	0.48533	0.4922	0.4915
Within R^2	0.32595	0.31629	0.19398	0.18928	0.18685

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A12: The Dynamics of the Shock

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked $\times \mathbb{1}(t-6)$	-0.241*	-0.241*	0.187	0.151	0.162
	(0.136)	(0.136)	(0.124)	(0.111)	(0.109)
Shocked $\times \mathbb{1}(t-5)$	-0.076	-0.076	0.145	0.156*	0.161*
	(0.124)	(0.124)	(0.096)	(0.090)	(0.092)
Shocked $\times \mathbb{1}(t-4)$	-0.237**	-0.237**	-0.046	-0.028	-0.016
	(0.109)	(0.109)	(0.074)	(0.066)	(0.065)
Shocked $\times \mathbb{1}(t-3)$	-0.216*	-0.216*	-0.088	-0.073	-0.047
	(0.126)	(0.126)	(0.094)	(0.090)	(0.084)
Shocked $\times \mathbb{1}(t-2)$	-0.081	-0.081	-0.070	-0.028	0.012
	(0.098)	(0.098)	(0.082)	(0.074)	(0.073)
Shocked $\times \mathbb{1}(t-1)$	-0.183**	-0.183**	-0.112*	-0.073	-0.052
	(0.081)	(0.081)	(0.060)	(0.057)	(0.055)
Shocked $\times \mathbb{1}(t)$	0.259**	0.259**	0.280***	0.221***	0.220***
	(0.107)	(0.107)	(0.085)	(0.068)	(0.068)
Shocked $\times \mathbb{1}(t+1)$	0.205*	0.205*	0.191**	0.150**	0.122*
	(0.107)	(0.107)	(0.078)	(0.065)	(0.063)
Shocked $\times \mathbb{1}(t+2)$	0.489***	0.489***	0.402***	0.364***	0.331***
	(0.147)	(0.147)	(0.117)	(0.102)	(0.100)
Shocked $\times \mathbb{1}(t+3)$	0.540***	0.540***	0.368***	0.377***	0.342***
	(0.137)	(0.137)	(0.115)	(0.100)	(0.096)
Shocked $\times \mathbb{1}(t+4)$	0.752***	0.752***	0.554***	0.482***	0.444***
	(0.147)	(0.147)	(0.132)	(0.112)	(0.111)
Shocked $\times \mathbb{1}(t+5)$	0.929***	0.929***	0.648***	0.572***	0.525***
	(0.185)	(0.185)	(0.144)	(0.131)	(0.129)
Shocked $\times \mathbb{1}(t+6)$	1.03***	1.03***	0.587***	0.592***	0.544***
	(0.209)	(0.209)	(0.149)	(0.135)	(0.124)

Continued on next page

Table A12 – *Continued from previous page*

	(1)	(2)	(3)	(4)	(5)
Shocked$\times\mathbb{1}(t+7)$	0.872***	0.872***	0.408**	0.378***	0.346***
	(0.196)	(0.196)	(0.171)	(0.139)	(0.130)
Shocked$\times\mathbb{1}(t+8)$	0.760***	0.760***	0.266	0.260*	0.250*
	(0.200)	(0.200)	(0.172)	(0.152)	(0.144)
Shocked$\times\mathbb{1}(t+9)$	0.467***	0.467***	0.037	0.158	0.126
	(0.166)	(0.166)	(0.139)	(0.123)	(0.119)
Shocked$\times\mathbb{1}(t+10)$	0.683***	0.683***	0.273*	0.309**	0.239*
	(0.176)	(0.176)	(0.150)	(0.144)	(0.132)
Shocked$\times\mathbb{1}(t+11)$	-0.130	-0.130	-0.033	0.030	0.034
	(0.149)	(0.149)	(0.132)	(0.122)	(0.120)
Shocked$\times\mathbb{1}(t+12)$	-0.294***	-0.294***	0.068	0.080	0.088
	(0.111)	(0.111)	(0.095)	(0.089)	(0.087)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	746,956	746,956	746,956	746,956	746,956
R^2	0.48414	0.48414	0.58706	0.61761	0.62407
Within R^2	0.00153	0.00153	0.00107	0.00091	0.00072

Two-way (Year×Month & Issuer) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A13: Defaults

The table reports the results of the following regression: $\text{default}_{j,t\to t+12} = \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{default}_{j,t\to t+12}$ is a dummy variable equal to one when loan j defaults in the period between t and $t+12$, $\text{Shocked}_{j,t}$ is a dummy equal to one when the loan has been sold by distressed CLOs and $X_{j,t}$ is a matrix of fixed effects to control for loan characteristics. Column (1) includes year×month fixed effects; column (2) includes year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-0.023***	-0.009***	-0.000	-0.000	-0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Fixed-effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	141,564	141,564	141,564	141,564	141,564
R^2	0.052	0.129	0.408	0.471	0.483
Within R^2	0.00	0.000	0.000	0.000	0.000

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A14: Rating Changes

The table reports the results of the following regression: $y = \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where the outcome variable is either $y = \text{downgrade}_{j,t\to t+12}$ or $y = \text{upgrade}_{j,t\to t+12}$. $\text{downgrade}_{j,t\to t+12}$ is a dummy variable equal to one if loan j is downgraded between time t and $t+12$; $\text{upgrade}_{j,t\to t+12}$ is a dummy variable equal to one if loan j is upgraded between time t and $t+12$. $\text{Shocked}_{j,t}$ is a dummy equal to one when the loan has been sold by shocked CLOs and $X_{j,t}$ is a matrix of fixed effects to control for loan characteristics. Column (1) includes year×month fixed effects; column (2) adds year×month×time-to-maturity and year×month×rating fixed effects; column (3) adds year×month×industry and year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	$\text{downgrade}_{j,t\to t+12}$			**$\text{upgrade}_{j,t\to t+12}$**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Shocked	0.000	-0.000	-0.000	-0.000	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Fixed-effects						
Year×Month	Yes	Yes	Yes	Yes	Yes	Yes
Year×Month×TTM	No	Yes	Yes	No	Yes	Yes
Year×Month×Rating	No	Yes	Yes	No	Yes	Yes
Year×Month×Industry	No	No	Yes	No	No	Yes
Year×Month×Interest	No	No	Yes	No	No	Yes
Fit statistics						
Observations	75,489	75,484	75,405	75,489	75,484	75,405
R^2	0.222	0.270	0.278	0.171	0.262	0.268
Within R^2	0	0	0	0	0	0

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A15: Liquidity

The table reports the results of the following regression: $y = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{j,t} X_{j,t} \delta + \varepsilon_{j,t}$, where the outcome variable is either $y = \gamma_{j,t}$ or $y = \log(\text{Nr.Trades})$. $\gamma_{j,t}$ measures the liquidity of loans by using the covariance in price changes (Roll, 1984), while log(Nr.Trades) is the natural logarithm of the number of times a given loan has been traded by CLOs. $\text{Shocked}_{j,t}$ is a dummy equal to one when the loan has been sold by distressed CLOs and $\text{Post}_{j,t}$ is a dummy that is equal to one after a loan has been sold by a shocked CLO. $X_{j,t}$ is a matrix of fixed effects to control for loan characteristics.

	γ			**log(Nr.Trades)**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Shocked	0.561	0.446	0.387	-0.048***	0.054***	0.051***
	(0.399)	(0.349)	(0.333)	(0.014)	(0.012)	(0.013)
Shocked× Post	-0.293	-0.301	-0.290	0.039***	-0.015	-0.015
	(0.328)	(0.362)	(0.363)	(0.012)	(0.014)	(0.014)
Fixed-effects						
Year×Month	Yes	Yes	Yes	Yes	Yes	Yes
Year×Month×TTM	No	Yes	Yes	No	Yes	Yes
Year×Month×Rating	No	Yes	Yes	No	Yes	Yes
Year×Month×Industry	No	No	Yes	No	No	Yes
Year×Month×Interest	No	No	Yes	No	No	Yes
Fit statistics						
Observations	137,866	123,273	123,165	137,866	123,273	123,165
R^2	0.121	0.155	0.156	0.154	0.198	0.200
Within R^2	0.000	0.000	0.000	0.000	0.000	0.000

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A16: Par Building - Placebo Test: Downgrades to B3

The table reports the difference in par built between CLOs that have received a shock to the bucket of securities rated B3. Columns (1) and (2) report the results of the following regressions: $\text{gain}_{i,j,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,j,t} = 100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for purchases and $\text{gain}_{i,j,t} = -100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for sales; $\text{Constrained}_{i,t}$ is a dummy variable equal to one whenever the Junior (column (1)) or Senior (column(2)) slack of CLO i is between 0% and and 5% in period t; $\text{Shocked}_{i,t}$ is a dummy variable equal to one whenever the loans of CLO i have been downgraded to B3. Columns (3) and (4) report the results of the following regressions: $\text{gain}_{i,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,t} = 100 \times \left(\sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ and the other variables are defined as above. $\text{Constrained}_{i,t}$ refers to Junior tests in column (3) and to Senior tests in column (4). Standard errors are reported in parentheses and they are double clustered at the Year×Month & CLO Deal level.

	Individual Transactions		**Multiple Transactions**	
	(1)	**(2)**	**(3)**	**(4)**
(Intercept)	-0.004***	-0.002***	-0.078***	-0.051***
	(0.0004)	(0.0002)	(0.006)	(0.003)
Shocked	0.0005	0.0003*	0.005	0.004
	(0.0004)	(0.0002)	(0.009)	(0.003)
Constrained	0.002***	-0.008***	0.030***	-0.079***
	(0.0004)	(0.002)	(0.007)	(0.024)
Shocked × Constrained	-0.0003	0.002	-0.003	0.026
	(0.0005)	(0.002)	(0.010)	(0.028)
Fit statistics				
Observations	309,028	303,160	30,156	29,034
R^2	0.000	0.000	0.001	0.001
Adjusted R^2	0.000	0.000	0.000	0.001
OC Test	Junior	Senior	Junior	Senior

Two-way (Year×Month & CLO Deal) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A17: Price Pressure - Placebo Test: Downgrades to B3

The table reports the results of the following regression: $\text{discount}_{j,k,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times Post_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,k,t} = 100 \times \log(100/P_{j,k,t})$, $P_{j,k,t}$ is the price of loan j issued by firm k at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by CLOs that experienced downgrades to B3 is above median and their slack is between 0% and 5%, $\text{Post}_{j,t}$ is a dummy equal to one after loan j has received an above median selling volume by CLOs with downgrades to B3, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. All the regressions include the lagged average discount on the issuer computed as Avg. $\text{discount}_{k,t-1} = \frac{1}{J_k \times (t-1)} \sum_{j=1}^{J_k} \sum_{s=1}^{t-1} \text{discount}_{j,k,s}$, where J_k is the number of loans by issuer k actively traded. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	0.738***	0.597***	0.538***	0.427***	0.377***
	(0.176)	(0.151)	(0.122)	(0.111)	(0.107)
Shocked×Post	-0.204	-0.001	0.059	0.066	0.075
	(0.147)	(0.132)	(0.111)	(0.104)	(0.103)
Avg. Discount_{t-1}	0.853***	0.838***	0.709***	0.693***	0.690***
	(0.033)	(0.034)	(0.032)	(0.031)	(0.031)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	332,118	332,118	332,118	332,118	332,118
R^2	0.489	0.504	0.597	0.636	0.644
Within R^2	0.406	0.388	0.290	0.276	0.274

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A18: Probability of Subscribing a Loan

The table provides the results of the following regression in column (1): $\text{Subscribed}_{i,j,t} = \alpha + \beta \text{Previously Held}_{i,j,t} + \varepsilon_{i,j,t}$, where $\text{Subscribed}_{i,j,t}$ is dummy variable equal to one when a CLO deal i subscribed a loan issued by borrower j at time t and $\text{Previously Held}_{i,j,t}$ is a dummy variable equal to one when the CLO i has held a different loan from borrower j before time t. Column (2) reports the results of the following model: $\text{Subscribed}_{i,j,t} = \alpha_t + \beta \text{Previously Held}_{i,j,t} + \varepsilon_{i,j,t}$ where α_t are Year×Month fixed effects. Column(3) adds the age of a CLO Deal and the logarithm of total assets under management as controls: $\text{Subscribed}_{i,j,t} = \alpha_t + \beta \text{Previously Held}_{i,j,t} + \gamma_1 \text{Age}_{i,t} + \gamma_2 \text{Total Assets}_{i,t} + \varepsilon_{i,j,t}$. Column (4) adds CLO deals fixed effect interacted with the issuer's industry fixed effect $\alpha_{j,s}$: $\text{Subscribed}_{i,j,t} = \alpha_t + \alpha_{j,s} + \beta \text{Previously Held}_{i,j,t} + \gamma_1 \text{Age}_{i,t} + \gamma_2 \text{Total Assets}_{i,t} + \varepsilon_{i,j,t}$. Standard errors clustered at the Year×Month level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**
(Intercept)	0.033***			
	(0.001)			
Previous Held	0.098***	0.098***	0.121***	0.089***
	(0.006)	(0.006)	(0.008)	(0.008)
Age			-0.006***	-0.147***
			(0.0003)	(0.046)
Total Assets			0.023***	0.010***
			(0.001)	(0.002)
Fixed-effects				
Yesr×Month	No	Yes	Yes	Yes
CLO Deal×Industry	No	No	No	Yes
Fit statistics				
Observations	9,626,651	9,626,651	6,439,971	6,439,971
R^2	0.021	0.026	0.050	0.091
Within R^2	–	0.021	0.041	0.014

One-way (Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A19: All-in Spread Drawn

The table reports the results of the following regression: $\text{AISD}_{j,t} = \alpha_t + \alpha_j + \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, $\text{AISD}_{j,t}$ is the all-in drawn spread for issuer j at time t, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects; column (5) adds issuer fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and double clustered by Year×Month and Issuer.

	(1)	(2)	(3)	(4)	(5)
Shocked	0.550***	0.347***	0.232***	0.232***	0.084**
	(0.037)	(0.042)	(0.041)	(0.041)	(0.039)
Fixed-effects					
Year × Month	Yes	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Rating	No	No	No	Yes	Yes
Issuer	No	No	No	No	Yes
Fit statistics					
Observations	13,468	13,365	13,365	13,365	13,365
R^2	0.04806	0.06598	0.09166	0.09166	0.60933
Within R^2	0.00304	0.00121	0.00055	0.00055	7e-05

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A20: Fraction of Institutional Loans

The table reports the results of the following regression: Fraction Inst.$_{j,t} = \alpha_t + \alpha_j + \beta$Shocked$_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, Fraction Inst.$_{j,t}$ measures the number of institutional tranches as a fraction of the total number of tranches issued by issuer j at time t, Shocked$_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by distressed CLOs in the previous twelve months. Column (1) includes Year×Month fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects; column (5) adds issuer fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and double clustered by Year×Month and Issuer.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-0.117***	-0.082***	-0.080***	-0.055**	-0.039*
	(0.024)	(0.024)	(0.024)	(0.022)	(0.022)
Fixed-effects					
Year×Month	Yes	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Rating	No	No	No	Yes	Yes
Issuer	No	No	No	No	Yes
Fit statistics					
Observations	13,468	13,365	13,365	13,365	13,365
R^2	0.11799	0.20027	0.20301	0.23609	0.34322
Within R^2	0.00183	0.001	0.00095	0.00047	2e-04

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A21: Fraction of Dollars Borrowed

The table reports the results of the following regression: Fraction Inst. $\$_{j,t} = \alpha_t + \alpha_j + \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, Fraction Inst. $\$_{j,t}$ measures the total amount of dollars borrowed using institutional tranches as a fraction of the total amount borrowed by issuer j at time t, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects; column (5) adds issuer fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and double clustered by Year×Month and Issuer.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-0.113***	-0.099***	-0.092***	-0.059**	-0.055**
	(0.029)	(0.030)	(0.030)	(0.027)	(0.027)
Fixed-effects					
Year×Month	Yes	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Rating	No	No	No	Yes	Yes
Issuer	No	No	No	No	Yes
Fit statistics					
Observations	8,969	8,667	8,667	8,667	8,667
R^2	0.2085	0.23172	0.24261	0.31183	0.54704
Within R^2	0.00245	0.00195	0.00169	0.00077	0.00073

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A22: Institutional Tranches Size

The table reports the results of the following regression: $\log(\text{Inst. Tranche Size})_{j,t} = \alpha_t + \alpha_j + \beta \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\log(\text{Inst. Tranche Size})_{j,t}$ is the logarithm of the tranche size for institutional loans measured in dollars, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects; column (5) adds issuer fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and double clustered by Year×Month and Issuer.

	(1)	(2)	(3)	(4)	(5)
Shocked	-0.342***	-0.326***	-0.334***	-0.270***	-0.235***
	(0.111)	(0.109)	(0.106)	(0.096)	(0.077)
Fixed-effects					
Year×Month	Yes	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Rating	No	No	No	Yes	Yes
Issuer	No	No	No	No	Yes
Fit statistics					
Observations	5,723	5,623	5,623	5,623	5,623
R^2	0.08508	0.12548	0.15397	0.24187	0.59158
Within R^2	0.0033	0.00318	0.00341	0.00247	0.00204

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A23: All-in Spread Drawn Within Lead Agent

The table reports the results of the following regression: $\text{AISD}_{j,t} = \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, $\text{AISD}_{j,t}$ is the all-in drawn spread for issuer j at time t, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month×Lead Agent fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and clustered by Year×Month.

	(1)	**(2)**	**(3)**	**(4)**
Shocked	0.519***	0.467***	0.467***	0.436***
	(0.097)	(0.094)	(0.094)	(0.090)
Fixed-effects				
Year×Month×Lead Agent	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes
Rating	No	No	Yes	Yes
Industry	No	No	No	Yes
Fit statistics				
Observations	13,252	12,826	12,826	12,826
R^2	0.63895	0.63905	0.63905	0.64029
Within R^2	0.00064	0.00051	0.00051	0.00043

One-way (Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A24: Fraction of Institutional Loans Within Lead Agent

The table reports the results of the following regression: Fraction $\text{Inst.}_{j,t} = \beta \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, Fraction $\text{Inst.}_{j,t}$ measures the number of institutional tranches as a fraction of the total number of tranches issued by issuer j at time t, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by distressed CLOs in the previous twelve months. Column (1) includes Year×Month×Lead Agent fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and clustered by Year×Month.

	(1)	**(2)**	**(3)**	**(4)**
Shocked	-0.227**	-0.216***	-0.210***	-0.213***
	(0.091)	(0.078)	(0.076)	(0.079)
Fixed-effects				
Year×Month×Lead Agent	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes
Rating	No	No	Yes	Yes
Industry	No	No	No	Yes
Fit statistics				
Observations	13,252	12,826	12,826	12,826
R^2	0.60046	0.64498	0.64604	0.65359
Within R^2	0.00068	0.00069	0.00064	0.00067

One-way (Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A25: Fraction of Dollars Borrowed Within Lead Agent

The table reports the results of the following regression: Fraction Inst. $\$_{j,t} = \beta \text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, Fraction Inst. $\$_{j,t}$ measures the total amount of dollars borrowed using institutional tranches as a fraction of the total amount borrowed by issuer j at time t, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month×Lead Agent fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and clustered by Year×Month.

	(1)	**(2)**	**(3)**	**(4)**
Shocked	-0.169*	-0.183**	-0.182**	-0.180**
	(0.090)	(0.093)	(0.085)	(0.089)
Fixed-effects				
Year×Month×Lead Agent	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes
Rating	No	No	Yes	Yes
Industry	No	No	No	Yes
Fit statistics				
Observations	8,794	8,457	8,457	8,457
R^2	0.95913	0.96312	0.9645	0.96552
Within R^2	0.00533	0.00685	0.00693	0.00681

One-way (Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A26: Institutional Tranches Size Within Lead Agent

The table reports the results of the following regression: $\log(\text{Inst. Tranche Size})_{j,t} = \beta\text{Shocked}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\log(\text{Inst. Tranche Size})_{j,t}$ is the logarithm of the tranche size for institutional loans measured in dollars, $\text{Shocked}_{j,t}$ is a dummy variable that is equal to one when firm j has been sold by shocked CLOs in the previous twelve months. Column (1) includes Year×Month×Lead Agent fixed effects; column (2) adds time-to-maturity fixed effects constructed by grouping the variable in ten buckets; column (3) adds industry fixed effects; column (4) adds rating fixed effects. Rating is constructed from the closest rating available for the firm. Standard errors are reported in paretheses and clustered by Year×Month.

	(1)	**(2)**	**(3)**	**(4)**
Shocked	-0.156	-0.181	-0.155	-0.189
	(0.203)	(0.192)	(0.223)	(0.237)
Fixed-effects				
Year×Month×Lead Agent	Yes	Yes	Yes	Yes
TTM	No	Yes	Yes	Yes
Rating	No	No	Yes	Yes
Industry	No	No	No	Yes
Fit statistics				
Observations	5,727	5,622	5,622	5,622
R^2	0.82295	0.82543	0.82998	0.85369
Within R^2	0.00015	2e-04	0.00014	0.00024

One-way (Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.2 Figures

Figure A1: Fraction of Caa (CCC) securities

The upper panel reports a histogram of the fraction of Caa (CCC) rated securities, while the lower panel superposes a local polynomial approximation of the density with the relative confidence intervals, following Cattaneo et al. (2020).

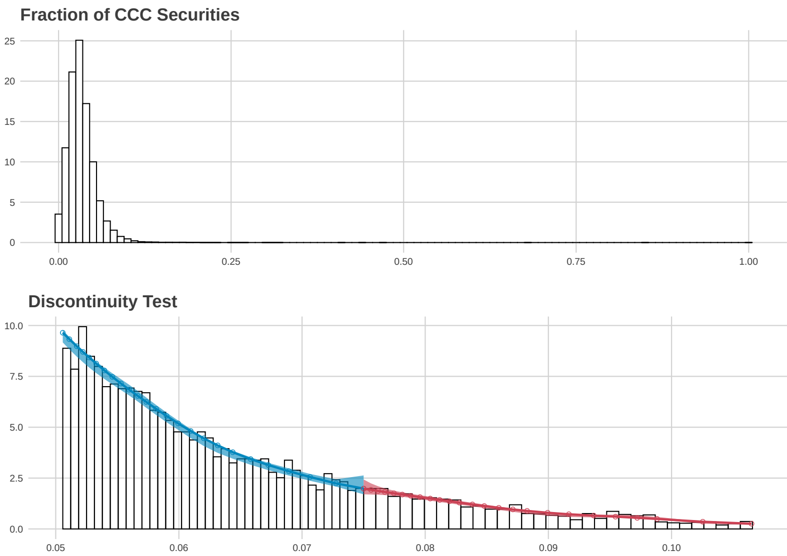


Figure A2: Par Building and OC Tests Slack

The figure depicts the gain in par as a function of the slack of the over-collateralization test by plotting the estimated coefficients of the following regression: $\text{gain}_{i,t} = \sum_{s=1}^{S} \beta_s \mathbb{1}_s + \varepsilon_{i,t}$, where $\text{gain}_{i,t} = 100 \times \left(\sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{PrincipalBalance_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{PrincipalBalance_{i,t}} \right)$; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior (left panel) or Senior (right panel) slack belongs to bucket s of the following $S = 7$ buckets: [-1.00,-0.05), [-0.05,0), [0,0.05), [0.05,0.10), [0.10,0.15), [0.15,0.20), [0.20,1.00). Full results are reported in Table A29.

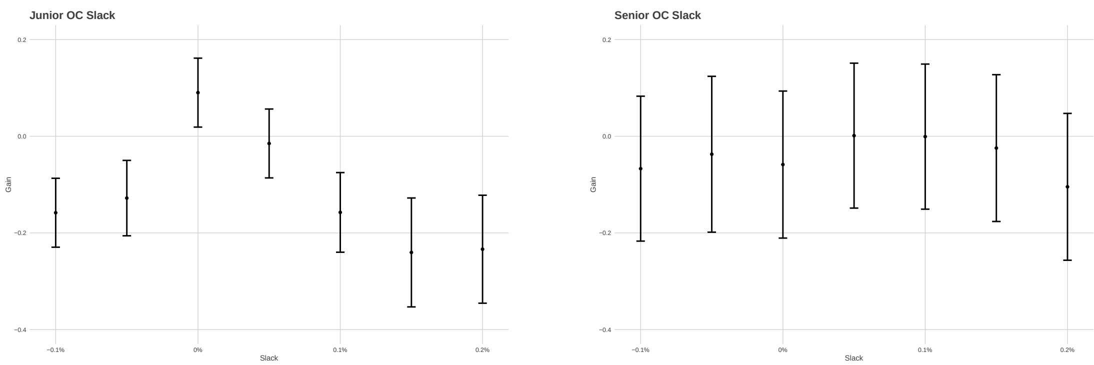


Figure A3: Fraction of Above-Par Securities Sold in the Month of a Downgrade to Caa (CCC)

The plots report the amount of securities sold above par as a fraction of total sales as a function of the slack in Junior OC tests. Observations are binned following Cattaneo et al. (2019). Each panel fits a separate polynomial of order $p = 1, 2, 3, 4$ to observation with positive and negative slack, following Calonico et al. (2015).

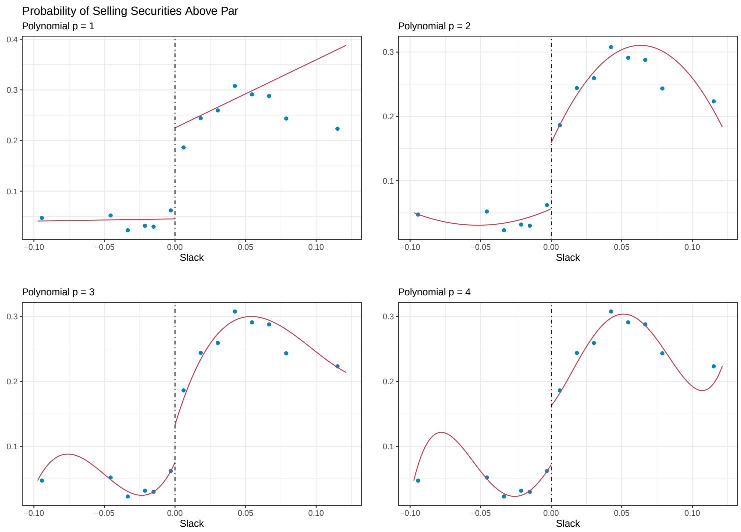


Figure A4: Fraction of Caa (CCC) rated Securities Sold in the Month of a Downgrade to Caa (CCC)

The plots report the amount of securities rated at Caa (CCC) or below sold as a fraction of total sales as a function of the slack in Junior OC tests. Observations are binned following Cattaneo et al. (2019). Each panel fits a separate polynomial of order $p = 1, 2, 3, 4$ to observation with positive and negative slack, following Calonico et al. (2015).

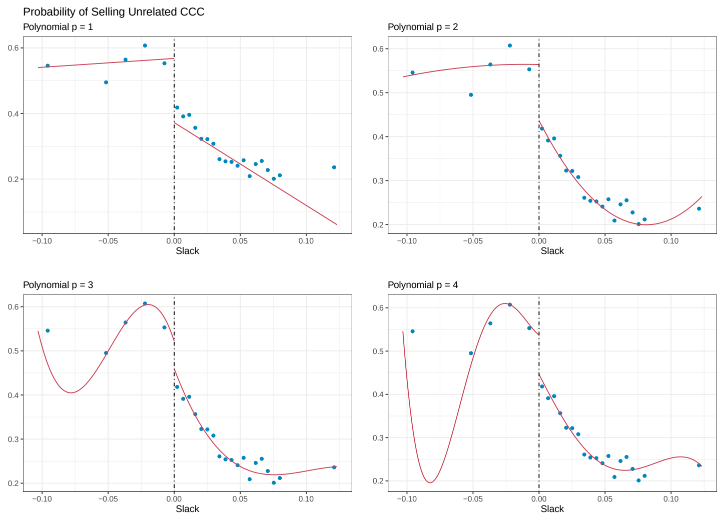


Figure A5: Price Pressure

The figure plots the results of the following regression: $\text{discount}_{j,t} = \gamma \text{Shocked}_{j,t} + \sum_{s=-6}^{12} \beta_s \text{Shocked}_{j,t} \times \mathbb{1}(t+s) + X_{j,t}\delta + \varepsilon_{j,t}$, where $discount_{j,t} = 100 \times \log(100/P_{j,t})$, $P_{j,t}$ is the price of loan j at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by shocked CLOs is above median, $\mathbb{1}(t+s)$ is a set of dummies that are equal to one $s = -6, -5, ..., 11, 12$ months around the event of the sale at time t, $X_{j,t}$ is a matrix containing the following fixed effects: year×yonth×time-to-maturity, year×month×rating, year×month×industry. Time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. Standard errors are two-way clustered at the year×month and issuer level.

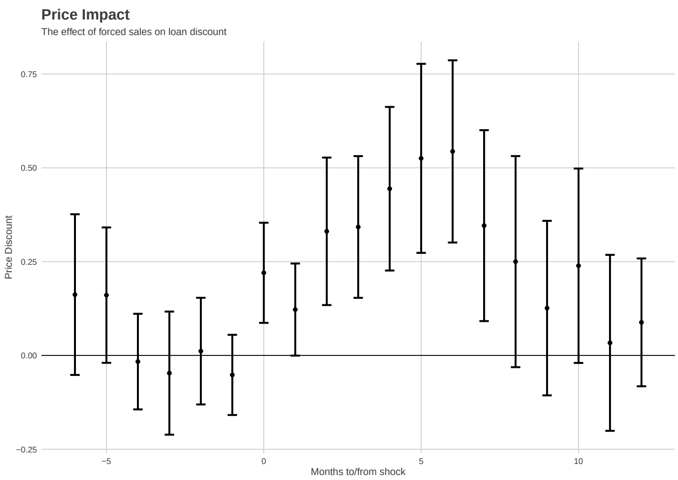
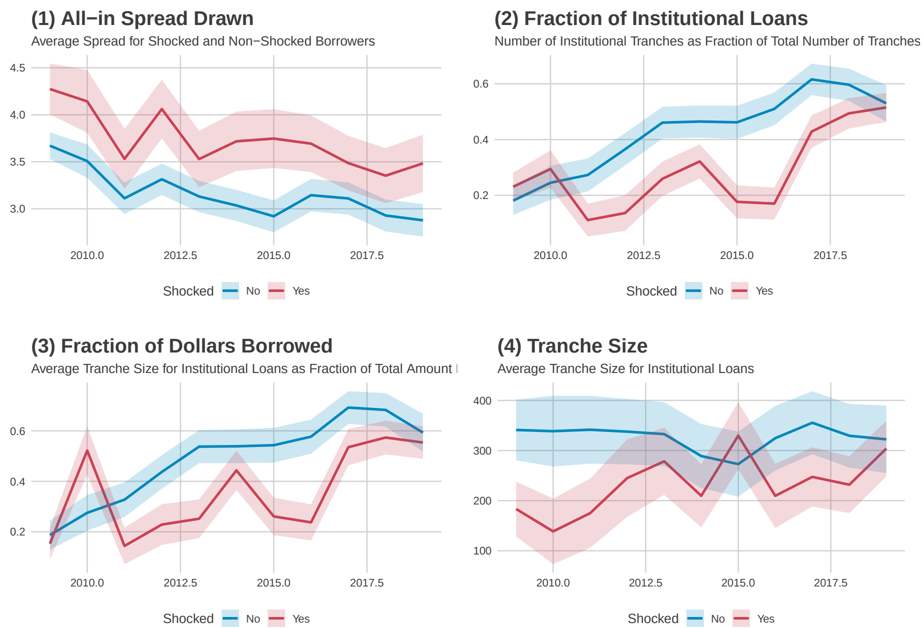


Figure A6: The Outcome in the Primary Market

The figure is composed of fours panels. (1) reports the yearly average all-in spread drawn in basis points; (2) reports the fraction of institutional loans per year as the total number of institutional tranches as a fraction of the total number of tranches; (3) reports the total amount of funds raised via institutional tranches as a fraction of total amount of funds raised; (4) reports the yearly average tranche size for institutional loans. All the panels reports the statistics for shocked firms in red and for the control group in blue.



A.3 Data Matching

This section describes the matching process to link loans in the CLO-i dataset with loans in the SDC Platinum dataset. Matching datasets on the U.S. syndicated loan market is a challenging task and no agreed upon procedure exists (Cohen et al., 2018). The task is even more challenging for the CLO-i dataset: due to the fact that information on loan holdings and transactions are obtained from CLOs' reports, there is no guarantee of consistency across time and deals. This implies that, in order to guarantee the integrity of the data, most of the matching procedure requires human discretion. In what follows I describe the steps I undertaken to match CLO-i with SDC Platinum, which - to the best of my knowledge - is the first attempt in these regards.

The procedure starts from the CLO holdings in the CLO-i dataset from which I keep only the following types: "Term Loan B", "Term Loan C", "Term Loan D", "Term Loan (Other)" and "Other". These types represent the greatest majority of CLO holdings as Figure A11 shows. For each of these securities I proceed to compute the first and last date they appear in sample together with the maturity date provided on the report; Figure A12 shows their distribution. I then proceed by recoding the maturity date as the maximum between the last time a security appears in sample and the reported maturity date, given that a security cannot appear on a balance sheet past its maturity and the resulting maturity date likely stems from a reporting mistake. I then proceed to compare the information at the security level across different CLO reports. Whenever a security presents different maturity dates I keep the modal date across reports.

Unfortunately, for the greatest part of the sampe, CLO-i does not provide a unique identifier for each security[1]. In order to make sure that the same security does not appear twice or more because of spelling mistakes I use the following convention: two securities are considered the same if the agree on the name of the borrower, type and maturity date. This implies that the analysis in the paper is carried out at the level of the loan issue-tranche type. In order to guarantee the

[1]CLO-i provides a unique loan identifier only after 2019.

integrity of the match, issuer and security names have been removed of irrelevant information: special characters have been removed, all the names have been converted into lower cased letters, commonly used abbreviations (e.g., "corp", "grp", etc...) have been fixed, similarly to Cohen et al. (2018). Once a security has been uniquely identified on CLO-i, I keep information regarding the first time it appears in sample, the maturity date, the first rating and the first interest rate, whenever they are available.

I then proceed to apply the same procedure to the universe of loans in SDC Platinum, which - however - does not require to be fixed for duplicates. First, I obtain the distinct observations by borrower, announcement date, facility and tranche. Then the borrower names are adjusted as previously described.

I then match the refined samples of loans in CLO-i and SDC Platinum using the following criteria: first, the announcement date in SDC Platinum needs to be no more than one year away from the first time the loan appears in CLO-i, second the maturity date in SDC Platinum is no more than one year away from the maturity date in the CLO-i, third at least one word in the borrower name reported in SDC Platinum needs to appear in either the issuer or the security name in CLO-i[2]. This leaves me with approximately 300,000 matches. These matches have been manually inspected in order to avoid false positives. When a security in CLO-i is matched with more than one security in SDC Platinum I keep the closest security according to the following criteria: first, I give priority to the securities sharing the same maturity date and loan type; second, I keep securities whose types are consistent[3]; third, if more than one tranche is matched, I keep the closest loan in terms of maturity and interest rate.

The final matching rate by year is displayed in Figure A25 for the securities in

[2]Applying fuzzy string matching routines provides a reduced matching rate compared to this procedure.

[3]CLO-i provides information on the loan type which I augment with information from the security name whenever possible. I adopt the following matching rules. "Term Loan B" is matched with the following SDC Platinum types: "1stLienTermLoan", "Other", ""RevCred/Term Ln", "RevCred/TLB", "Term Loan", and"Term Loan B". "Term Loan C" is matched with the following SDC Platinum types: "1stLienTermLoan", "Other", ""RevCred/Term Ln", "RevCred/TLC", "Term Loan","Term Loan C", and "Term Loan/LC". "Term Loan D" is matched with the following SDC Platinum types: "1stLienTermLoan", "Other", ""RevCred/Term Ln", "RevCred/TLD", "Term Loan","Term Loan D". "Term Loan (Other)" and "Other" are matched with all the previous.

the CLO-i dataset, and in Figure A26 for the securities in SDC Platinum.

A.4 Additional Tables

Table A27: Largest Deals and Management Teams

The table reports the twenty largest CLO managers, in Panel A, and CLO deal, in Panel B, based on the average total assets in sample. For managers, total assets are the sum of the current balance of holdings managed in a given month, while for CLO deals the sum of the current balance of holdings held in a given report. Assets are measured in $Mlns.

Panel A: Managers				
Name	Mean	Median	Min.	Max.
GSO Capital Partners	15536.34	15075.26	5813.30	27456.15
Carlyle Group	11932.33	10541.84	1620.47	24850.72
KKR Financial Advisors	11376.15	11419.91	9645.77	13058.09
Och Ziff	11136.66	11622.85	8459.28	13089.39
Credit Suisse Asset Management	11093.56	8881.74	2718.49	28504.00
Ares Management	7819.85	7326.49	1999.78	19644.97
Barings	7701.79	7383.15	3856.02	17866.87
PGIM	7606.96	5128.54	1653.70	24309.48
Alcentra	7457.40	7629.65	2525.83	13718.95
CIFC Asset Management	7011.29	6284.61	493.14	18429.52
KKR	6994.95	6402.08	2700.69	18720.42
Apollo Global Management	6918.01	6946.41	1303.09	14200.04
CBAM	6581.15	7750.09	1231.18	10223.07
CVC Credit Partners	6205.31	5197.08	2134.08	13792.22
Barclays Capital	6151.88	5772.62	4755.84	9752.23
MJX Asset Management	5584.94	4169.99	852.10	17248.10
Investcorp Credit Management	5510.20	5238.23	2053.56	11178.54
Sculptor	5082.09	5269.00	588.07	14447.15
Octagon Credit Investors	4859.36	2993.79	670.60	18102.80
Onex Credit Partners	4848.65	3993.32	501.63	10095.17
Panel B: CLO Deals				
Name	Mean	Median	Min.	Max.
Heta Funding	5331.63	4938.22	4721.55	7897.35
GoldenTree Credit Opportunities Financing I	3652.07	3652.07	3652.07	3652.07
Fortress Credit Opportunities I	3441.43	3454.84	1251.67	18336.84
KKR Financial CLO 2007-1	2817.67	3196.07	56.26	3424.96
GSO Domestic Capital Funding	2101.24	2101.24	170.94	4031.53
Antares CLO 2017-1	2100.21	2111.52	2016.51	2171.98
Prospect Funding I	1817.85	1431.03	1189.79	3750.56
CBAM 2017-2	1552.42	1554.13	1532.76	1572.30
RR 3	1479.54	1489.09	1306.71	1523.67
Ares XXXI-R	1362.82	1266.47	1243.70	2546.44
Genesis CLO 2007-2	1324.87	1313.84	1275.03	1371.04
Fortress Credit Opportunities IX	1318.64	1407.09	1108.52	1491.66
CBAM 2017-3	1288.33	1286.42	1262.60	1309.47
Tennenbaum Opportunities Partners V	1246.03	1331.87	623.67	3839.90
CBAM 2017-1	1239.99	1240.39	1225.05	1257.22
Ares XXXI	1213.49	1228.08	1075.40	1266.95
Antares CLO 2017-2	1208.99	1215.78	1164.59	1241.41
Churchill Financial Cayman	1198.26	1176.02	192.14	5162.72
Woodmont 2017-2 Trust	1171.74	1171.65	1103.75	1210.65
Zohar II 2005-1	1137.45	1153.06	919.02	1202.65

Table A28: Holding vs. Market Prices - All Transactions

The table compares the discount of loans as they are reported by CLOs with the closest market price in the following year by running the following regression: $\text{discount}_{j,t} = \alpha_j + \alpha_t + \beta_1 \text{Transaction}_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$ is the discount at which a loan is recorded or traded compared to par, $\text{Transaction}_{j,t}$ is an indicator variable equal to one whenever the price comes from an actual transaction (sale or purchase) and zero otherwise, α_j and α_t are issuer and time fixed effects, while $X_{j,t}$ includes a set of fixed effects for rating, industry and interest rate of the loan. Panel A consider the universe of matched loans, while Panel B focuses on loans rated Caa (CCC).

	Panel A: All Loans				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Transaction	0.700***	0.709***	0.709***	0.709***	0.708***
	(0.160)	(0.161)	(0.161)	(0.161)	(0.161)
Fit statistics					
Observations	2,598,448	2,580,708	2,580,708	2,580,708	2,598,448
R^2	0.570	0.602	0.603	0.603	0.729
Within R^2	0.003	0.003	0.003	0.003	0.005
	Panel B: Caa (CCC) Loans				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Transaction	2.96***	2.97***	2.97***	2.97***	3.01***
	(0.669)	(0.669)	(0.669)	(0.669)	(0.670)
Fit statistics					
Observations	228,440	228,440	228,440	228,440	228,440
R^2	0.628	0.648	0.650	0.650	0.739
Within R^2	0.0083	0.00884	0.0089	0.0089	0.01218
Fixed-Effects					
Year×Month	Yes	Yes	Yes	Yes	No
Issuer	Yes	Yes	Yes	Yes	No
Rating	No	Yes	Yes	Yes	Yes
Industry	No	No	Yes	Yes	Yes
Interest	No	No	No	Yes	Yes
Year×Month×Issuer	No	No	No	No	Yes

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A29: Par Building and OC Test Slack

Columns (1) and (2) report the results of the following regressions: $\text{gain}_{i,j,t} = \sum_{s=1}^{S} \mathbb{1}_s + \varepsilon_{i,t}$, where $\text{gain}_{i,j,t} = (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}}$ for purchases and $\text{gain}_{i,j,t} = -(100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}}$ for sales; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior (column (1)) or Senior (column(2)) slack belongs to bucket s of the following $S = 7$ buckets: [-1.00,-0.05), [-0.05,0), [0,0.05), [0.05,0.10), [0.10,0.15), [0.15,0.20), [0.20,1.00). Columns (3) and (4) report the results of the following regressions: $\text{gain}_{i,t} = \sum_{s=1}^{S} \mathbb{1}_s + \varepsilon_{i,t}$, where $\text{gain}_{i,t} = \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{PrincipalBalance_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{PrincipalBalance_{i,t}}$; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior (column (3)) or Senior (column(4)) slack belongs to bucket s.

	Individual Transactions		**Multiple Transactions**	
	(1)	**(2)**	**(3)**	**(4)**
$\mathbb{1}[-1.00, -0.05)$	-0.037***	-0.008	-0.158***	-0.067
	(0.012)	(0.012)	(0.043)	(0.091)
$\mathbb{1}[-0.05, 0.00)$	0.012	0.0004	0.030	0.030
	(0.012)	(0.014)	(0.048)	(0.098)
$\mathbb{1}[0.00, 0.05)$	0.041***	0.004	0.249***	0.008
	(0.012)	(0.012)	(0.043)	(0.093)
$\mathbb{1}[0.05, 0.10)$	0.036***	0.009	0.143***	0.068
	(0.012)	(0.012)	(0.043)	(0.091)
$\mathbb{1}[0.10, 0.15)$	0.014	0.009	0.0007	0.066
	(0.012)	(0.012)	(0.050)	(0.092)
$\mathbb{1}[0.15, 0.20)$	0.001	0.006	-0.082	0.043
	(0.014)	(0.012)	(0.069)	(0.093)
$\mathbb{1}[0.20, 1.00)$	0.013	-0.006	-0.075	-0.038
	(0.014)	(0.012)	(0.068)	(0.093)
Fit statistics				
Observations	301,770	301,669	18,539	17,676
R^2	0.01193	0.00368	0.04887	0.01655
Adjusted R^2	0.01191	0.00366	0.04856	0.01621
OC Test	Junior	Senior	Junior	Senior

Two-way (CLO Deal & Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A30: Rating Factor and OC Test Slack

The table presents the average rating factor for loans sold, in column (1), and purchased, column (2), by CLOs as a function of the slack of their Junior OC tests, by reporting the coefficients of the following regression: $\text{RF}_{i,j,t} = \sum_{s=1}^{S} \beta_s \mathbb{1}_s + \varepsilon_{i,j,t}$; where $\text{RF}_{i,j,t}$ is the rating factor of loan j, sold by CLO i at time t; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior slack belongs to bucket s of the following $S = 7$ buckets: [-1.00,-0.05), [-0.05,0), [0,0.05), [0.05,0.10), [0.10,0.15), [0.15,0.20), [0.20,1.00). Column (3) aggregates the results by reporting the coefficients of the following regression: $\Delta\text{WARF}_{i,t} = \sum_{s=1}^{S} \beta_s \mathbb{1}_s + \varepsilon_{i,t}$, where $\Delta\text{WARF}_{i,t} = \sum_j \text{RF}_{i,j,t} \times \frac{\text{Amt. Purchased}_{i,j,t}}{\sum_j \text{Amt. Purchased}_{i,j,t}} - \sum_j \text{RF}_{i,j,t} \times \frac{\text{Amt. Sold}_{i,j,t}}{\sum_j \text{Amt. Sold}_{i,j,t}}$. Standard errors clustered by CLO Deal & Year×Month are reported in parentheses.

	Individual Transactions		**Multiple Transactions**
	(1)	**(2)**	**(3)**
$\mathbb{1}[-1.00, -0.05)$	4392.7***	2669.5***	-397.4
	(597.4)	(78.5)	(467.8)
$\mathbb{1}[-0.05, 0.00)$	46.1	54.2	79.2
	(631.8)	(98.2)	(475.6)
$\mathbb{1}[0.00, 0.05)$	-1951.6***	138*	751.1
	(597.7)	(80.5)	(468.3)
$\mathbb{1}[0.05, 0.10)$	-1576.2***	39.1	-59.9
	(597.2)	(80.8)	(469.3)
$\mathbb{1}[0.10, 0.15)$	-1135.7*	84.6	-736.6
	(628.1)	(211.6)	(538.5)
$\mathbb{1}[0.15, 0.20)$	-459	501.9	-1840.4**
	(854.1)	(390.5)	(762.2)
$\mathbb{1}[0.20, 1.00)$	-1617.9***	5.06	-438.1
	(601)	(100.7)	(559.2)
Fit statistics			
Observations	155,079	162,629	21,043
R^2	0.05856	0.00231	0.1135
Adjusted R^2	0.05843	0.00222	0.11295

Two-way (CLO Deal & Year×Month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A31: Price Pressure

The table reports the results of the following regression: $\text{discount}_{j,t} = \beta_1 \text{Shocked}_{i,j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{i,j,t} + X_{j,t}\delta + \varepsilon_{i,j,t}$, where $\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$, $P_{j,t}$ is the price of loan j at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by shocked CLOs is above median, $\text{Post}_{i,j,t}$ is a dummy equal to one after loan j has received an above median selling volume by shocked CLOs, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes issuer and year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-0.520***	-0.144*	-0.069	-0.073	-0.067
	(0.095)	(0.078)	(0.056)	(0.051)	(0.054)
Shocked×Post	1.02***	0.690***	0.525***	0.514***	0.499***
	(0.144)	(0.121)	(0.091)	(0.086)	(0.089)
Fixed-Effects					
Issuer	Yes	Yes	Yes	Yes	Yes
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	738,354	738,354	738,354	738,354	738,354
R^2	0.421	0.432	0.533	0.564	0.540
Within R^2	0.324	0.312	0.224	0.218	0.223

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A32: Price Pressure - Purchases

The table reports the results of the following regression: $\text{discount}_{j,t} = \beta_1 \text{Shocked}_{i,j,t} + \beta_2 \text{Shocked}_{j,t} \times \text{Post}_{i,j,t} + X_{j,t}\delta + \varepsilon_{i,j,t}$, where $\text{discount}_{j,t} = 100 \times \log(100/P_{j,t})$, $P_{j,t}$ is the price of loan j at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by shocked CLOs is above median, $\text{Post}_{i,j,t}$ is a dummy equal to one after loan j has received an above median selling volume by shocked CLOs, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes issuer and year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	-1.4***	-1.03***	-0.582***	-0.581***	-0.558***
	(0.140)	(0.121)	(0.089)	(0.089)	(0.086)
Shocked×Post	-0.068	-0.338***	-0.199	-0.199	-0.207
	(0.110)	(0.121)	(0.131)	(0.131)	(0.134)
Fixed-Effects					
Issuer	Yes	Yes	Yes	Yes	Yes
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	746,956	746,956	600,004	600,004	598,787
R^2	0.48686	0.48926	0.56682	0.56815	0.5695
Within R^2	0.0068	0.00438	0.00173	0.00172	0.00165

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A33: The Dynamics of the Shock - Purchases

	(1)	(2)	(3)	(4)	(5)
Shocked$\times\mathbb{1}(t-6)$	-0.122	-0.122	-0.042	-0.095	-0.090
	(0.114)	(0.114)	(0.099)	(0.086)	(0.084)
Shocked$\times\mathbb{1}(t-5)$	-0.228**	-0.228**	-0.121	-0.108	-0.069
	(0.103)	(0.103)	(0.088)	(0.082)	(0.080)
Shocked$\times\mathbb{1}(t-4)$	-0.144	-0.144	-0.081	-0.128*	-0.072
	(0.088)	(0.088)	(0.084)	(0.072)	(0.073)
Shocked$\times\mathbb{1}(t-3)$	-0.106	-0.106	-0.017	-0.014	0.025
	(0.126)	(0.126)	(0.108)	(0.109)	(0.106)
Shocked$\times\mathbb{1}(t-2)$	-0.164	-0.164	-0.037	-0.021	0.033
	(0.146)	(0.146)	(0.137)	(0.128)	(0.126)
Shocked$\times\mathbb{1}(t-1)$	-0.188*	-0.188*	-0.108	-0.096	-0.027
	(0.105)	(0.105)	(0.112)	(0.110)	(0.107)
Shocked$\times\mathbb{1}(t)$	-0.342***	-0.342***	-0.229**	-0.197*	-0.120
	(0.111)	(0.111)	(0.112)	(0.108)	(0.105)
Shocked$\times\mathbb{1}(t+1)$	-0.466***	-0.466***	-0.322***	-0.305***	-0.228**
	(0.100)	(0.100)	(0.103)	(0.103)	(0.102)
Shocked$\times\mathbb{1}(t+2)$	-0.502***	-0.502***	-0.334***	-0.299***	-0.233**
	(0.107)	(0.107)	(0.111)	(0.107)	(0.104)
Shocked$\times\mathbb{1}(t+3)$	-0.445***	-0.445***	-0.277**	-0.245**	-0.198*
	(0.109)	(0.109)	(0.112)	(0.108)	(0.107)
Shocked$\times\mathbb{1}(t+4)$	-0.492***	-0.492***	-0.333***	-0.309***	-0.251**
	(0.102)	(0.102)	(0.104)	(0.101)	(0.099)
Shocked$\times\mathbb{1}(t+5)$	-0.269**	-0.269**	-0.117	-0.171	-0.162
	(0.128)	(0.128)	(0.128)	(0.109)	(0.103)
Shocked$\times\mathbb{1}(t+6)$	-0.254***	-0.254***	-0.090	-0.127	-0.102
	(0.092)	(0.092)	(0.102)	(0.100)	(0.101)

Continued on next page

Table A33 – *Continued from previous page*

	(1)	(2)	(3)	(4)	(5)
Shocked $\times \mathbb{1}(t+7)$	-0.121	-0.121	-0.0010	-0.042	-0.036
	(0.108)	(0.108)	(0.116)	(0.107)	(0.109)
Shocked $\times \mathbb{1}(t+8)$	-0.212**	-0.212**	-0.069	-0.088	-0.084
	(0.106)	(0.106)	(0.118)	(0.116)	(0.117)
Shocked $\times \mathbb{1}(t+9)$	-0.148	-0.148	-0.047	-0.045	-0.043
	(0.099)	(0.099)	(0.106)	(0.105)	(0.106)
Shocked $\times \mathbb{1}(t+10)$	-0.140	-0.140	-0.040	-0.061	-0.054
	(0.098)	(0.098)	(0.105)	(0.102)	(0.102)
Shocked $\times \mathbb{1}(t+11)$	0.070	0.070	0.154	0.148	0.124
	(0.120)	(0.120)	(0.123)	(0.119)	(0.120)
Shocked $\times \mathbb{1}(t+12)$	0.020	0.020	0.075	0.070	0.055
	(0.097)	(0.097)	(0.088)	(0.081)	(0.081)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	435,712	435,712	435,712	435,712	435,712
R^2	0.5226	0.5226	0.57876	0.61972	0.63059
Within R^2	0.00238	0.00238	0.00152	0.00129	0.00096

Two-way (Year×Month & Issuer) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A34: Probability of Default Conditional on Having a Defaulted Loan

The table reports the results of the following regressions: $y = \alpha + \beta \text{default}_{i,t} + \varepsilon_{i,j,t}$. The outcome variable is $y = \text{default}_{i,j,t}$ or $y = \text{default}_{i,j,t \to t+12}$. $\text{default}_{i,j,t}$ is a dummy variable equal to one if loan j from issuer i is in default at time t; $\text{default}_{i,j,t \to t+12}$ is a dummy variable equal to one if loan j from issuer i defaults between time t and $t+12$. The independent variable is $\text{default}_{i,t}$, a dummy variable equal to one when any of the loans of issuer i are in default at time t. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses. The sample contains only loans from issuers with more than one issue currently on CLOs portfolios.

	$\text{default}_{\mathbf{i,j,t}}$	$\text{default}_{\mathbf{i,j,t \to t+12}}$
(Intercept)	0.000	0.000
	(0.000)	(0.000)
$\text{default}_{i,t}$	0.200***	0.285***
	(0.012)	(0.014)
Fit statistics		
Observations	2,548,697	2,548,697
R^2	0.15203	0.22163
Adjusted R^2	0.15203	0.22163

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A35: Par Building - Placebo Test: Downgrades to B2

The table reports the difference in par built between CLOs that have received a shock to the bucket of securities rated B2. Columns (1) and (2) report the results of the following regressions: $\text{gain}_{i,j,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,j,t} = 100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for purchases and $\text{gain}_{i,j,t} = -100 \times \left((100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ for sales; $\text{Constrained}_{i,t}$ is a dummy variable equal to one whenever the Junior (column (1)) or Senior (column(2)) slack of CLO i is between 0% and and 5% in period t; $\text{Shocked}_{i,t}$ is a dummy variable equal to one whenever the loans of CLO i have been downgraded to B2. Columns (3) and (4) report the results of the following regressions: $\text{gain}_{i,t} = \alpha + \beta_1 \text{Constrained}_{i,t} + \beta_2 \text{Shocked}_{i,t} + \beta_3 \text{Constrained}_{i,t} \times \text{Shocked}_{i,t} + \varepsilon_{i,t}$, where $\text{gain}_{i,t} = 100 \times \left(\sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans bought}_{i,j,t}}{\text{Principal Balance}_{i,t}} - \sum_j (100 - P_{j,t-1}) \times \frac{\text{Nr. loans sold}_{i,j,t}}{\text{Principal Balance}_{i,t}} \right)$ and the other variables are defined as above. $\text{Constrained}_{i,t}$ refers to Junior tests in column (3) and to Senior tests in column (4). Standard errors are reported in parentheses and they are double clustered at the Year×Month & CLO Deal level.

	Individual Transactions		**Multiple Transactions**	
	(1)	**(2)**	**(3)**	**(4)**
(Intercept)	-0.004***	-0.002***	-0.080***	-0.048***
	(0.0004)	(0.0002)	(0.009)	(0.003)
Shocked	0.0005	0.000	0.010	-0.0009
	(0.0004)	(0.0001)	(0.009)	(0.003)
Constrained	0.002***	-0.006***	0.034***	-0.065**
	(0.0004)	(0.002)	(0.009)	(0.027)
Shocked × Constrained	-0.0005	-0.0001	-0.011	-0.001
	(0.0004)	(0.002)	(0.009)	(0.028)
Fit statistics				
Observations	309,028	303,160	30,156	29,034
R^2	0.000	0.000	0.001	0.001
Adjusted R^2	0.000	0.000	0.000	0.001
OC Test	Junior	Senior	Junior	Senior

Two-way (Year×Month & CLO Deal) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A36: Price Pressure - Placebo Test: Downgrades to B2

The table reports the results of the following regression: $\text{discount}_{j,k,t} = \beta_1 \text{Shocked}_{j,t} + \beta_2 \text{Shocked}_{j,t} \times Post_{j,t} + X_{j,t}\delta + \varepsilon_{j,t}$, where $\text{discount}_{j,k,t} = 100 \times \log(100/P_{j,k,t})$, $P_{j,k,t}$ is the price of loan j issued by firm k at time t, $\text{Shocked}_{j,t}$is a dummy variable equal to one when loan j selling volume by CLOs that experienced downgrades to B2 is above median and their slack is between 0% and 5%, $\text{Post}_{j,t}$ is a dummy equal to one after loan j has received an above median selling volume by CLOs with downgrades to B2, $X_{j,t}$ is a matrix containing various fixed effects and controls. Column (1) includes year×month fixed effects; column (2) adds year×month×time-to-maturity fixed effects; column (3) adds year×month×rating fixed effects; column (4) adds year×month×industry fixed effects; column (5) adds year×month×interest rate fixed effects. Interest and time-to-maturity fixed effects are constructed after bucketing the continuous variable in ten groups. All the regressions include the lagged average discount on the issuer computed as $\text{Avg. discount}_{k,t-1} = \frac{1}{J_k \times (t-1)} \sum_{j=1}^{J_k} \sum_{s=1}^{t-1} \text{discount}_{j,k,s}$, where J_k is the number of loans by issuer k actively traded. Two-way clustered standard errors at the year×month and issuer level are reported in parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Shocked	0.713***	0.565***	0.506***	0.403***	0.335***
	(0.172)	(0.145)	(0.112)	(0.106)	(0.106)
Shocked×Post	-0.180	0.039	0.083	0.088	0.109
	(0.135)	(0.117)	(0.098)	(0.093)	(0.092)
Avg. Discount_{t-1}	0.852***	0.838***	0.709***	0.693***	0.690***
	(0.033)	(0.034)	(0.032)	(0.031)	(0.031)
Fixed-Effects					
Year×Month	Yes	No	No	No	No
Year×Month×TTM	No	Yes	Yes	Yes	Yes
Year×Month×Rating	No	No	Yes	Yes	Yes
Year×Month×Industry	No	No	No	Yes	Yes
Year×Month×Interest	No	No	No	No	Yes
Fit statistics					
Observations	332,118	332,118	332,118	332,118	332,118
R^2	0.489	0.504	0.597	0.636	0.644
Within R^2	0.406	0.388	0.290	0.276	0.274

Two-way (Year×Month & Issuer) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

A.5 Additional Figures

Figure A7: Fraction of Caa (CCC) and Defaulted Securities - Time Series

The upper plot displays the times series of the median fraction of securities rated Caa (CCC) or below. The lower plot displays the times series of the median fraction of defaulted securities. Shaded areas indicate the 25th and 75th percentiles of the distribution.

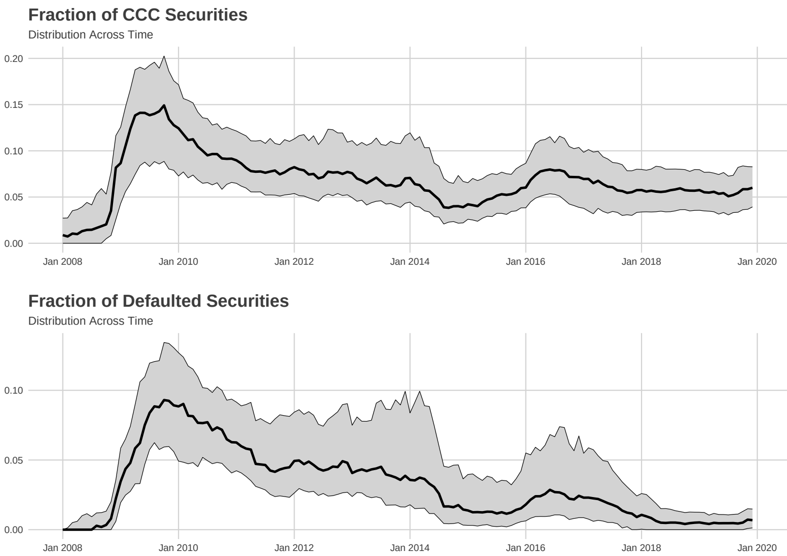


Figure A8: Fraction of Caa (CCC) and Defaulted Securities by CLO Deal's Age

The plots report the median fraction of securities rated Caa (CCC) or below, on the left, and the median fraction of defaulted securities, on the right, as a function of a CLO deal's age. The upper plots measure age as a fraction of the total age of each deal, while the bottom plots measure age in years.Shaded areas indicate the 25th and 75th percentiles of the distribution.

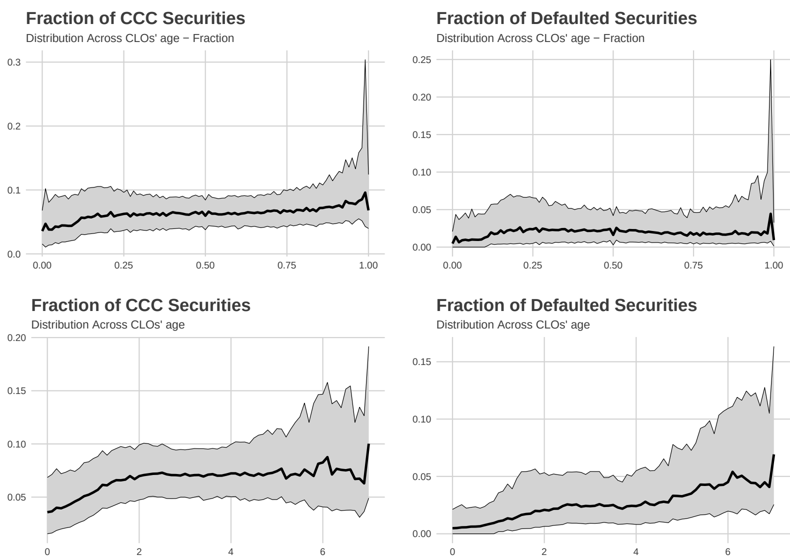


Figure A9: Loan Prices

The upper plot reports the time series of the median price of loans traded by CLOs by month, while the lower plot refers to the average price weighted by the volume of trades. Blue lines include all the transactions by CLOs, red lines include only loans purchased by CLOs while green lines include only loans sold by CLOs.

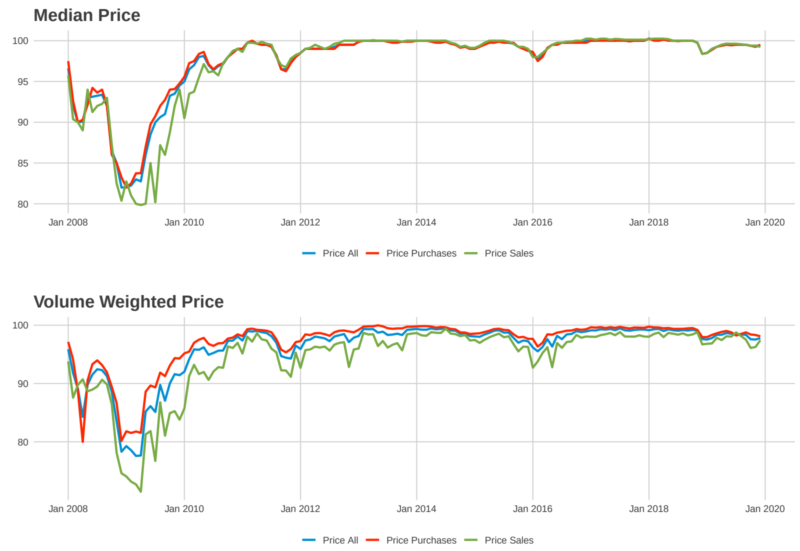


Figure A10: Fraction of Variation in Discounts Explained

The plots reports the fraction of variation in discounts explained by various characteristics as the R^2 of a regression of discounts on various fixed effects, i.e. $discount_{j,t} = X_{j,t}\beta + \varepsilon_{j,t}$, where $discount_{j,t} = 100 \times \log\left(100/P_{j,t}\right)$, $P_{j,t}$ is the price of loan j at time t and $X_{j,t}$ is a matrix of fixed effects. (1) includes Year $\times$ Month fixed effects; (2) interested rate fixed effects interacted with Year $\times$ Month fixed effects, where the interest rate of a loan is grouped in ten buckets; (3) includes time-to-maturity fixed effects interacted with Year $\times$ Month fixed effects, where time-to-maturity is grouped in ten buckets; (4) includes industry fixed effects interacted with Year $\times$ Month fixed effects; (5) includes rating fixed effects interacted with Year $\times$ Month fixed effects; (6) includes loan issuer fixed effects; (7) includes all the previous fixed effects; (8) includes issuer $\times$ Year $\times$ Month fixed effects; (9) includes issuer $\times$ Year $\times$ Month fixed effects and, separately, interest, time-to-maturity, industry and rating fixed effects interacted with Year $\times$ Month fixed effects.

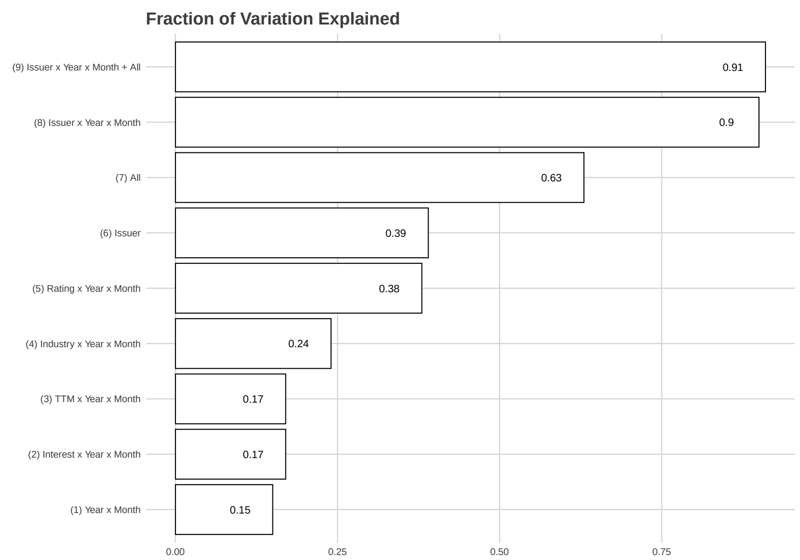


Figure A11: Securities Held by CLOs by Type

The figure plots a histogram of the securities held by CLOs bucketed by their type. The upper plot counts the number of securities, while the bottom plot refers to each security has a fraction of CLO assets.

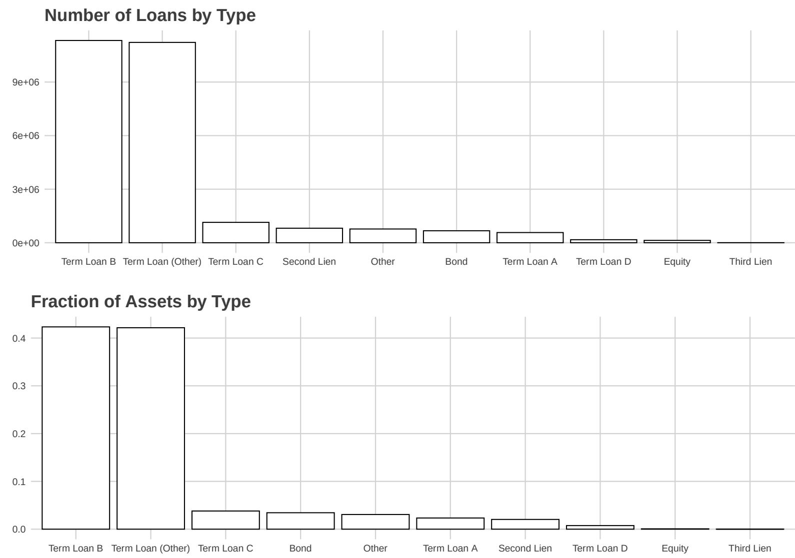


Figure A12: Securities Held by CLOs

The topmost figure plots a histogram of the securities held by CLOs bucketed by the first time they appear in sample. The middle histogram refers to the last date each security appears in sample. The bottom histogram refers to the maturity date of each security.

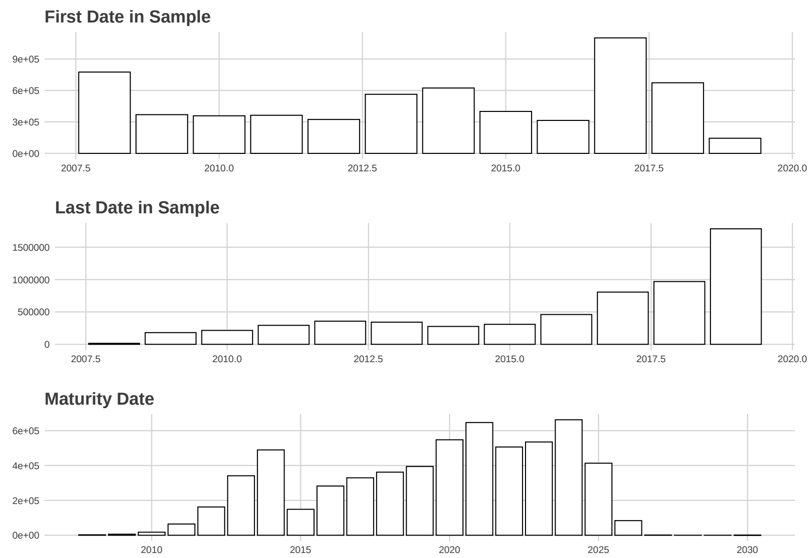


Figure A13: OC Tests Slack

The plot reports the monthly time series of the median slack for over-collateralization (OC) tests together with their 25th and 75th percentiles. Senior OC tests are in red, while Junior OC tests are in blue. For each deal the slack of tranche k is constructed as $\text{slack}_k = \frac{\text{test result}_k - \text{test threshold}_k}{\text{test threshold}_k}$. The senior slack for each deal is obtained as the median slack of tranches A and B, while the junior slack is obtained as the median slack of the remaining tranches.

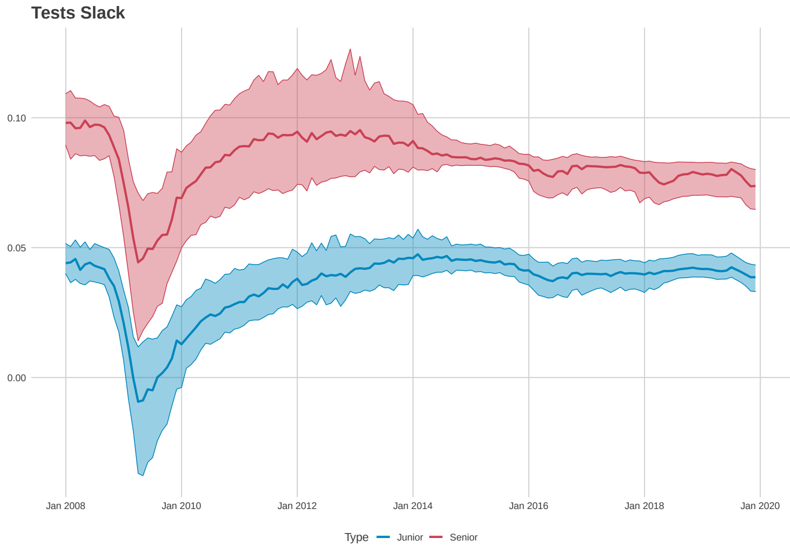


Figure A14: Weighted-Average Rating Factor

The plot reports the time series of the median weighted-average rating factor, in red, and its average, in blue. Both statistics are computed from the cross-section of CLO deals reporting in any given month.

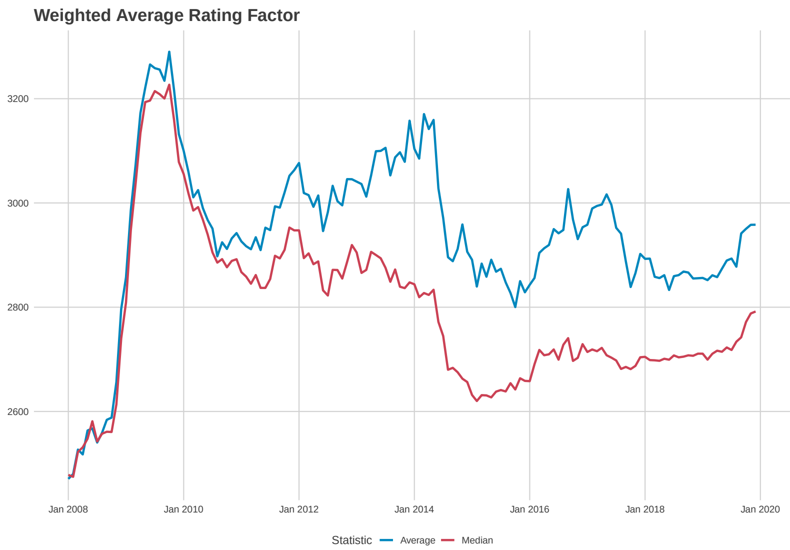


Figure A15: Number of Transactions Per Loan

The upper panel displays the average number of transactions per loan per month. The bottom panel displays the median number of transactions per loan per month. Blue lines represent loans purchased by CLOs, while red lines loans sold by CLOs.

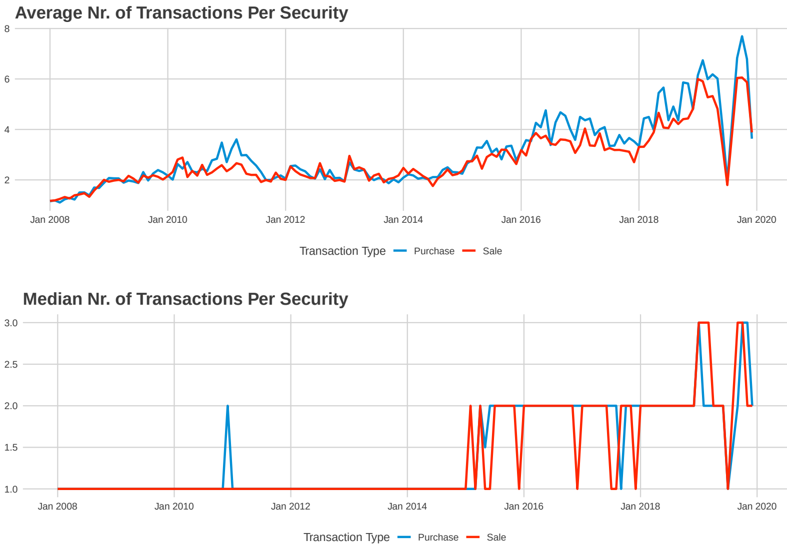


Figure A16: Par Gained and OC Test Slack

The plots report the gain in par as a function of the slack in Junior OC tests. Observations are binned following Cattaneo et al. (2019). Each panel fits a separate polynomial of order $p = 1, 2, 3, 4$ to observation with positive and negative slack, following Calonico et al. (2015).

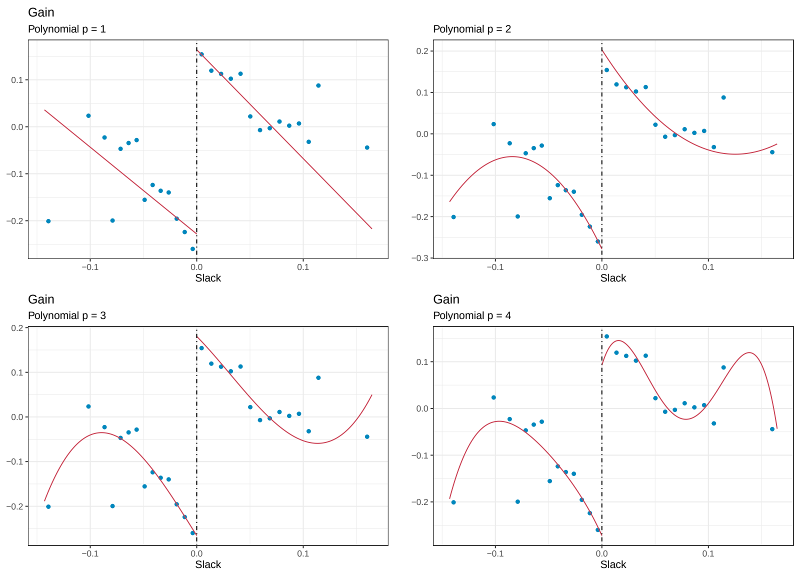


Figure A17: Rating Factor and OC Test Slack

The plots show the average rating factor for loans sold (left) and purchased (right) by CLOs as a function of the slack of their Junior OC tests, by reporting the coefficients of the following regression: $RF_{i,j,t} = \sum_{s=1}^{S} \beta_s \mathbb{1}_s + \varepsilon_{i,j,t}$; where $RF_{i,j,t}$ is the rating factor of loan j, sold by CLO i at time t; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior slack belongs to bucket s of the following $S = 7$ buckets: [-1.00,-0.05), [-0.05,0), [0,0.05), [0.05,0.10), [0.10,0.15), [0.15,0.20), [0.20,1.00).

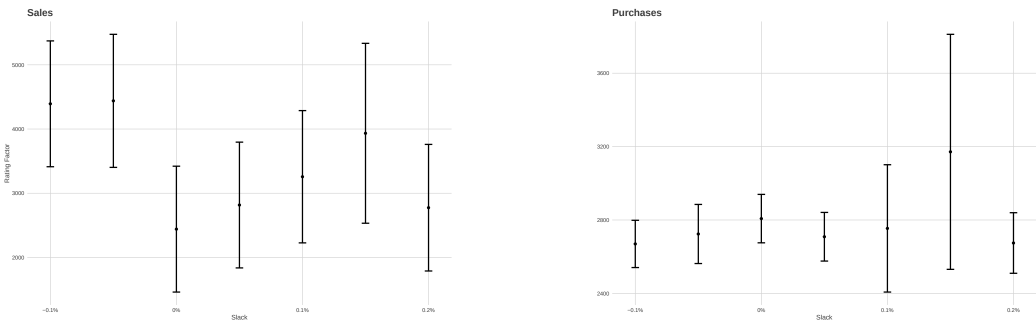


Figure A18: Rating Factor and OC Test Slack

The plot reports the coefficients of the following regression: $\Delta WARF_{i,t} = \sum_{s=1}^{S} \beta_s \mathbb{1}_s + \varepsilon_{i,t}$, where $\Delta WARF_{i,t} = \sum_j RF_{i,j,t} \times \frac{\text{Amt. Purchased}_{i,j,t}}{\sum_j \text{Amt. Purchased}_{i,j,t}} - \sum_j RF_{i,j,t} \times \frac{\text{Amt. Sold}_{i,j,t}}{\sum_j \text{Amt. Sold}_{i,j,t}}$; $RF_{i,j,t}$ is the rating factor of loan j, sold by CLO i at time t; $\mathbb{1}_s$ is a dummy variable equal to one whenever the Junior slack belongs to bucket s of the following $S = 7$ buckets: [-1.00,-0.05), [-0.05,0), [0,0.05), [0.05,0.10), [0.10,0.15), [0.15,0.20), [0.20,1.00).

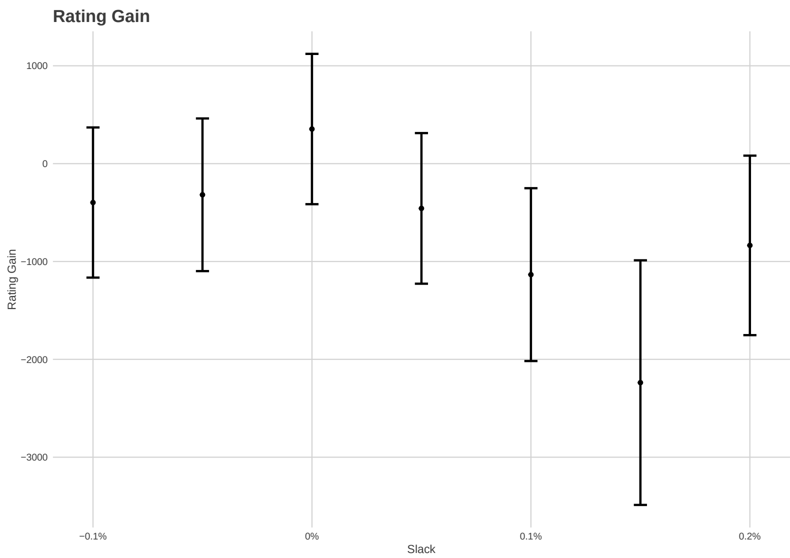


Figure A19: Probability of Selling Securities Above Par Around a Downgrade to Caa (CCC)

The plot reports the difference between CLOs that just passed their OC test and those that failed them in the fraction of above-par securities sold as a fraction of total sales.

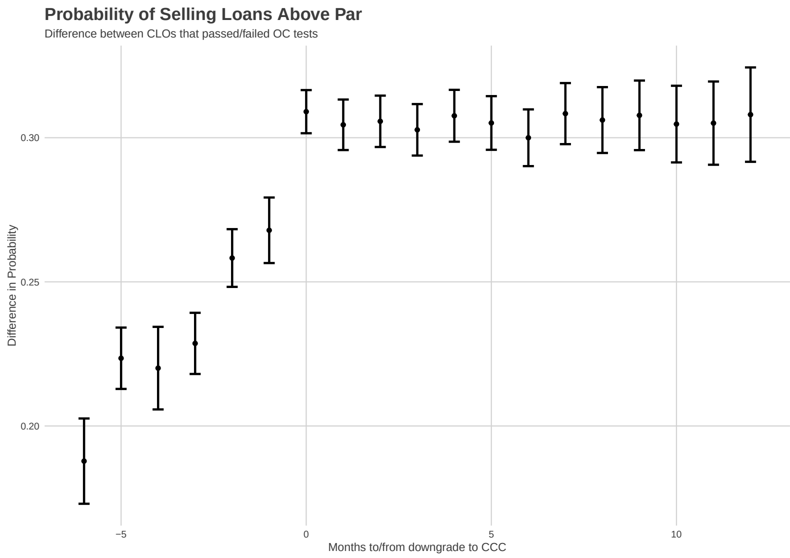


Figure A20: Probability of Selling Caa (CCC) Rated Securities Around a Downgrade to Caa (CCC)

The plot reports the difference between CLOs that just passed their OC test and those that failed them in the fraction of securities rated Caa (CCC) or below sold as a fraction of total sales.

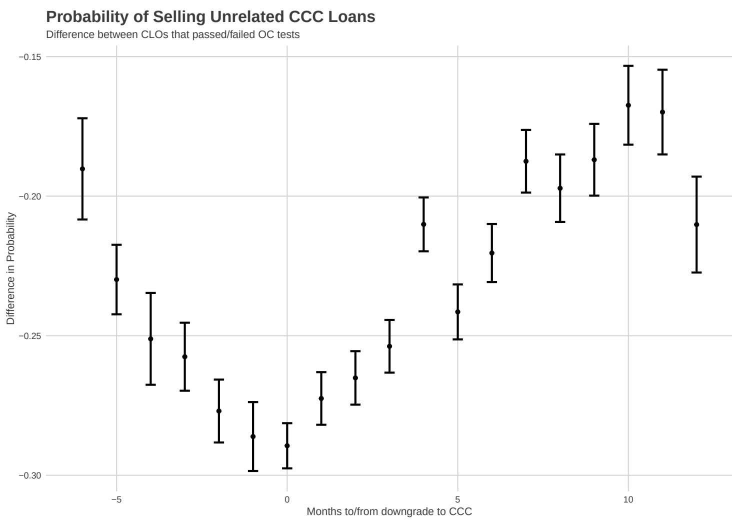


Figure A21: The Dynamics of the Shock - Sales

The figure plots the coefficients of models (1), (2), (3) and (5) in Table A12.

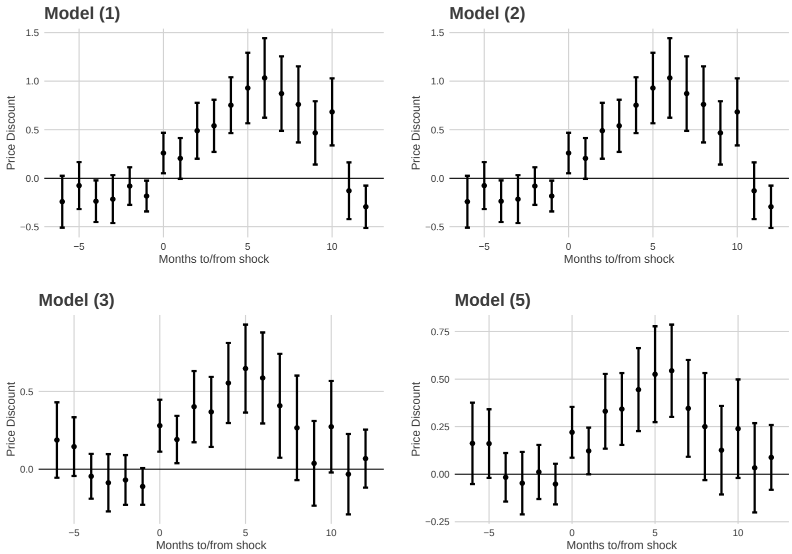


Figure A22: Average Discount for the Treated Loans

The plots reports the average discount for treated loans by around the month they are sold by shocked CLOs obtained from the following regression: $\text{discount}_{j,t} = \sum_{s=-6}^{12} \beta_s \mathbb{1}(t+s) + \varepsilon_{j,t}$, where $\text{discount}_{j,t} = 100 \times \log\left(100/P_{j,t}\right)$, $P_{j,t}$ is the price of loan j at time t and and $\mathbb{1}(t+s)$ is a set of dummies that are equal to one $s = -6, 5, ..., 11, 12$ months around the event of the sale at time t. Error bars reports the two-standard errors confidence intervals. Standard errors are two-wat clustered at the year×month and issuer level.

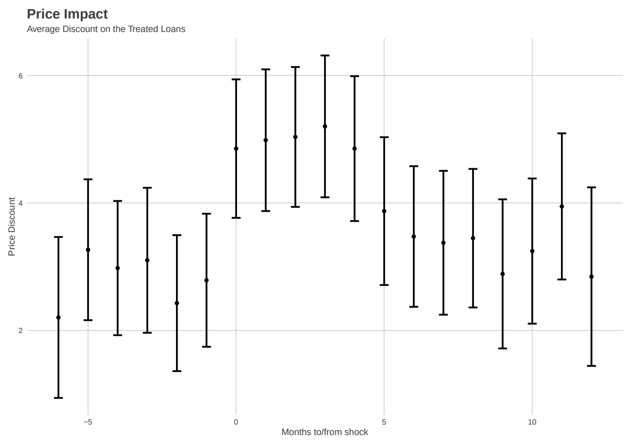


Figure A23: The Dynamics of the Shock - Purchases

The figure plots the coefficients of models (1), (2), (3), (4) and (5) from Table A33.

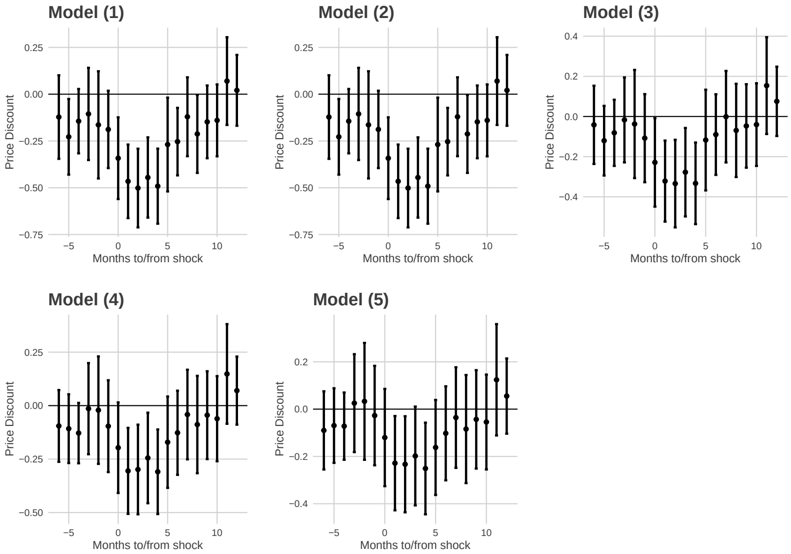


Figure A24: Loan Primary Use of Proceeds

The plots reports a histogram of the count of loans in the SDC Platinum dataset grouped by the primary use of proceeds. General includes general corporate purposes and payment of fees and expenses; Acquisition includes acquisition finance, future acquisitions, real estate and property acquisition and acquisition of securities; Refinancing includes general refinancing, bank refinancing, payments on previous borrowed money, payment on long-term borrowings, and down payments of previously borrowed money; LBO refers to leveraged buyouts; Lev. Recap. to any recapitalization; Proj. Finance to general project finance, recourse and non-recourse project finance; SBO refers to sponsored buyouts; Standby to standby facilities; Others contains all the remaining residual categories.

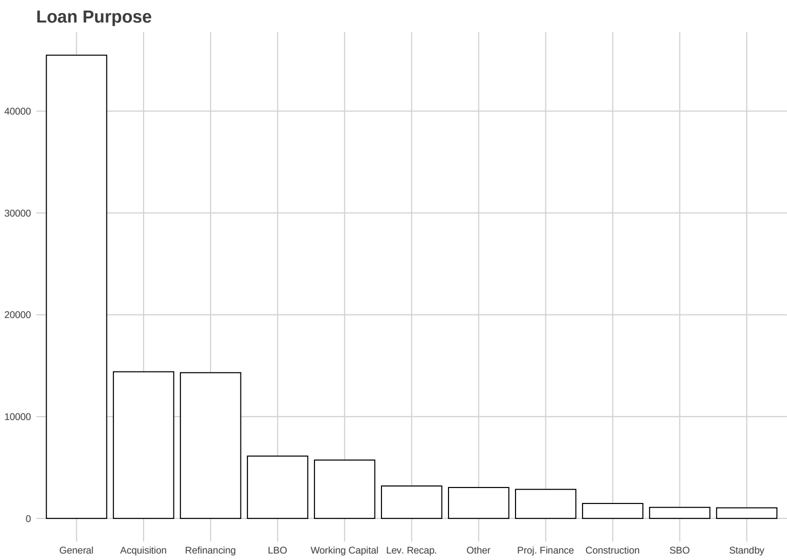


Figure A25: Assets Under Management and Number of Facilities Matched in CLO-i

The upper panel reports the amount of assets under management in the CLO-i dataset measured in $ Billions. The lower panel reports the number of facilities by year. Blue lines refer to the full sample, while red lines to the sample of securities matched with loans in the SDC Platinum dataset.

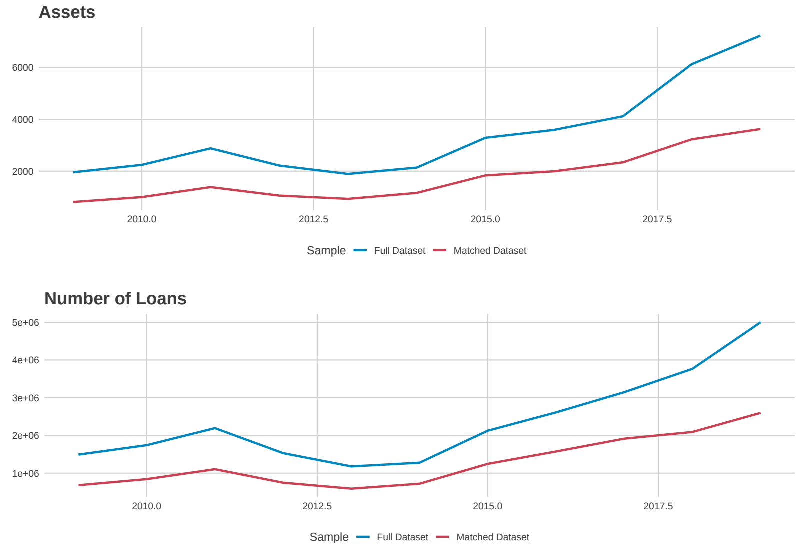
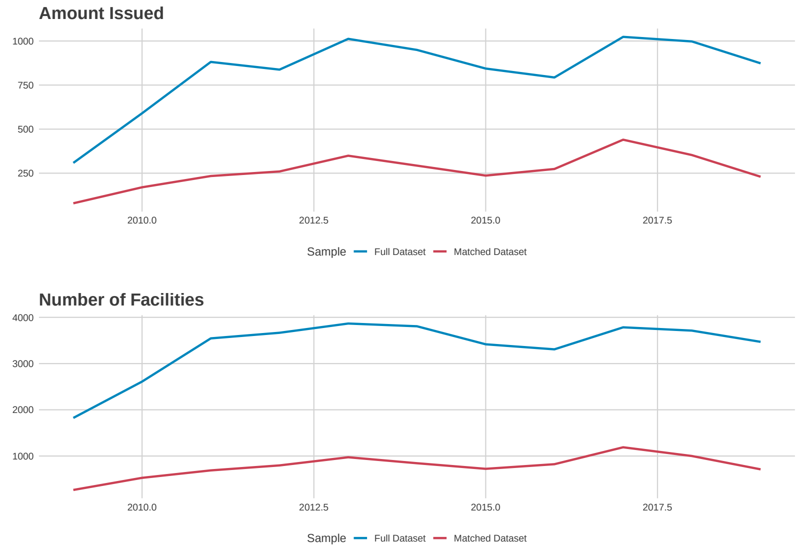


Figure A26: Amount and Number of Facilities Matched in SDC Platinum

The upper panel reports the amount of loans issued by year in the SDC Platinum dataset measured in $ Billions. The lower panel reports the number of facilities issued by year. Blue lines refer to the full sample, while red lines to the sample of leveraged loans matched with loans in the CLO-i dataset.



B. Appendix to Revealed Expectations and Learning Biases

B.1 Tables

Table B1: Summary Statistics

The table reports summary statistics for the data used. Column $\bar{x}$ reports the sample average of each variable, column σ its standard deviation, Min the smallest observation, Q1 the first quartile, Median the $50th$ percentile, Q3 the third quartile, Max the largest observation and N the number of observations. The first panel reports summary statistics regarding average and median past returns experienced by managers. The second panel reports six measures of expected excess returns computed as $\hat{\Sigma}_t \boldsymbol{w_{i,t}}$. Rows (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; rows (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Rows (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, rows (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and rows (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$. The third panel reports summary statistics on managers' careers; experience refers to the number of quarters since the first time a certain stock appeared in the manager's portfolio; max.experience refers to the maximum experience achieved for each manager-stock pair; tenure refers to the number of quarters since the first time the manager appeared in sample.

	$\bar{x}$	σ	Min	Q1	Median	Q3	Max	N
				Experienced Returns				
average	0.024	0.100	-0.557	-0.010	0.026	0.063	0.607	13,912,677
median	0.014	0.111	-0.871	-0.026	0.021	0.062	1.198	13,912,677
				Expected Excess Returns				
(1)	0.012	0.015	-0.282	0.004	0.007	0.014	1.336	5,416,032
(2)	0.011	0.014	-0.208	0.003	0.006	0.012	0.806	5,416,032
(3)	0.011	0.015	-0.161	0.003	0.006	0.013	0.764	5,416,032
(4)	0.012	0.015	-0.278	0.004	0.007	0.014	0.766	12,707,119
(5)	0.011	0.015	-0.292	0.003	0.006	0.012	1.086	12,707,119
(6)	0.011	0.015	-0.319	0.003	0.006	0.013	1.034	12,707,119
				Managers Careers				
experience	13.158	12.853	1	4	9	17	139	13,912,677
max. experience	13.884	11.981	1	6	11	17	139	1,223,610
tenure	26.896	21.943	1	10	21	39	139	75,179

Table B2: The Effect of Average Experienced Returns

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the standardised equal-weighted average experienced return, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.103***	0.103***	0.105***	0.149***	0.148***	0.151***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	1,270,823	1,270,823	1,270,823	2,856,830	2,856,830	2,856,830
R^2	0.781	0.765	0.773	0.709	0.692	0.695
Within-R^2	0.006	0.006	0.006	0.009	0.009	0.009
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B3: The Effect of Experienced Returns - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t \in \Delta T_{i,j,t}^q} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T_{i,j,t}^q}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.287***	0.272***	0.275***	0.273***	0.281***
	(0.008)	(0.013)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.134***	0.132***	0.136***	0.134***	0.132***	0.136***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)
β_3	0.041***	0.043***	0.040***	0.042***	0.042***	0.046***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
β_4	0.073***	0.073***	0.077***	0.075***	0.072***	0.078***
	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
β_5	0.238***	0.237***	0.241***	0.238***	0.237***	0.237***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
N	796,021	796,021	796,021	1,958,072	1,958,072	1,958,072
R^2	0.798	0.786	0.792	0.720	0.705	0.708
Within-R^2	0.042	0.043	0.043	0.043	0.042	0.042
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B4: The Effect of Experienced Returns - Ten Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.293***	0.258***	0.271***	0.290***	0.268***
	(0.012)	(0.039)	(0.010)	(0.010)	(0.017)	(0.007)
β_2	0.147***	0.163***	0.141***	0.149***	0.157***	0.148***
	(0.008)	(0.023)	(0.009)	(0.006)	(0.009)	(0.005)
β_3	0.100***	0.102***	0.098***	0.100***	0.102***	0.096***
	(0.006)	(0.011)	(0.006)	(0.004)	(0.005)	(0.004)
β_4	0.060***	0.058***	0.067***	0.059***	0.066***	0.061***
	(0.006)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)
β_5	0.028***	0.023***	0.021***	0.029***	0.030***	0.025***
	(0.005)	(0.008)	(0.005)	(0.003)	(0.003)	(0.003)
β_6	0.022***	0.024***	0.019***	0.021***	0.027***	0.024***
	(0.004)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)
β_7	0.020***	0.027***	0.026***	0.024***	0.020***	0.023***
	(0.004)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)
β_8	0.043***	0.045***	0.040***	0.045***	0.046***	0.046***
	(0.004)	(0.006)	(0.005)	(0.002)	(0.003)	(0.003)
β_9	0.080***	0.088***	0.087***	0.086***	0.088***	0.087***
	(0.006)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)
β_{10}	0.206***	0.204***	0.206***	0.208***	0.216***	0.215***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
N	442,353	442,353	442,353	1,073,779	1,073,779	1,073,779
R^2	0.824	0.812	0.820	0.750	0.736	0.738
Within-R^2	0.039	0.041	0.039	0.039	0.042	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B5: The Effect of Experienced Returns by Number of Managers

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Each column reports the results for the sub-sample of managers working in a team of $n_{i,t} \in \{1,2,3,4 \text{ or more}\}$, members at time t. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. The first four columns report results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the last four columns report results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.276***	0.114***	0.006**	0.015***	0.275***	0.114***	0.006***	0.005***
	(0.008)	(0.014)	(0.003)	(0.003)	(0.006)	(0.008)	(0.002)	(0.002)
β_2	0.133***	0.053***	0.004**	0.008***	0.134***	0.047***	0.004***	0.001
	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
β_3	0.040***	0.011***	0.004**	0.006***	0.041***	0.010***	0.006***	0.003**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
β_4	0.072***	0.014***	0.000	0.001	0.074***	0.015***	0.003***	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
β_5	0.237***	0.017***	0.002**	0.001	0.237***	0.019***	0.004***	0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	796,021	580,367	1,000,968	790,078	1,958,072	1,455,284	2,773,180	2,181,406
R^2	0.798	0.912	0.991	0.989	0.720	0.866	0.984	0.978
Within-R^2	0.042	0.002	0.000	0.001	0.043	0.003	0.001	0.000
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B6: Managers Who Have Switched Funds - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, 3, 4, 5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.224***	0.272***	0.209***	0.240***	0.214***	0.250***
	(0.018)	(0.031)	(0.023)	(0.014)	(0.014)	(0.022)
β_2	0.133***	0.116***	0.112***	0.125***	0.117***	0.124***
	(0.015)	(0.016)	(0.015)	(0.009)	(0.009)	(0.011)
β_3	0.048***	0.046***	0.031***	0.063***	0.049***	0.066***
	(0.011)	(0.014)	(0.011)	(0.008)	(0.007)	(0.009)
β_4	0.066***	0.071***	0.051***	0.078***	0.065***	0.073***
	(0.010)	(0.010)	(0.011)	(0.007)	(0.006)	(0.007)
β_5	0.199***	0.199***	0.202***	0.216***	0.219***	0.211***
	(0.011)	(0.013)	(0.013)	(0.007)	(0.007)	(0.009)
N	110,037	110,037	110,037	225,676	225,676	225,676
R^2	0.892	0.885	0.889	0.843	0.834	0.842
Within-R^2	0.034	0.038	0.034	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B7: Managers Who Have Switched Funds - Ten Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t \in \Delta T_{i,j,t}^q} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T_{i,j,t}^q}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.246***	0.253***	0.252***	0.281***	0.212***	0.198***
	(0.026)	(0.026)	(0.027)	(0.055)	(0.017)	(0.021)
β_2	0.129***	0.148***	0.127***	0.154***	0.124***	0.116***
	(0.018)	(0.018)	(0.020)	(0.023)	(0.011)	(0.013)
β_3	0.071***	0.102***	0.106***	0.101***	0.097***	0.071***
	(0.015)	(0.014)	(0.018)	(0.011)	(0.010)	(0.010)
β_4	0.054***	0.066***	0.074***	0.060***	0.055***	0.042***
	(0.015)	(0.012)	(0.017)	(0.009)	(0.008)	(0.009)
β_5	0.027**	0.035***	0.030***	0.040***	0.029***	0.022***
	(0.013)	(0.011)	(0.012)	(0.008)	(0.007)	(0.008)
β_6	0.026**	0.025**	0.015	0.029***	0.012*	0.023***
	(0.011)	(0.011)	(0.013)	(0.007)	(0.007)	(0.006)
β_7	0.011	0.019*	0.013	0.027***	0.020***	0.027***
	(0.011)	(0.010)	(0.011)	(0.006)	(0.007)	(0.007)
β_8	0.044***	0.031***	0.033**	0.056***	0.040***	0.038***
	(0.012)	(0.011)	(0.013)	(0.007)	(0.006)	(0.007)
β_9	0.090***	0.073***	0.085***	0.084***	0.086***	0.077***
	(0.012)	(0.013)	(0.012)	(0.008)	(0.007)	(0.007)
β_{10}	0.183***	0.169***	0.180***	0.195***	0.193***	0.200***
	(0.014)	(0.014)	(0.016)	(0.010)	(0.009)	(0.009)
N	78,920	78,920	78,920	160,237	160,237	160,237
R^2	0.914	0.915	0.914	0.869	0.865	0.867
Within-R^2	0.038	0.037	0.039	0.044	0.040	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B8: Learning Parameters

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k}\right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $r_{j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. Clustered standard errors are in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.146***	0.139***	0.144***	0.205***	0.205***	0.207***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
λ_1	-1.901***	-1.838***	-1.873***	-1.663***	-1.700***	-1.683***
	(0.068)	(0.064)	(0.064)	(0.034)	(0.038)	(0.035)
λ_2	-1.659***	-1.487***	-1.563***	-1.574***	-1.610***	-1.590***
	(0.108)	(0.116)	(0.108)	(0.053)	(0.061)	(0.053)
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B9: Risk Aversion - Pooled Regressions

The table reports the parameter estimates obtained from the following pooled regression: $r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w}^*_{\mathbf{i,t}})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. α is the pooled estimated bias across managers and time, γ is the pooled estimated risk aversion across managers and time. Standard errors are clustered at the manager-time and stock-time level and reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
α	0.011***	0.011***	0.011***	0.010***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
γ	0.915***	0.999***	0.958***	1.204***	1.283***	1.255***
	(0.079)	(0.082)	(0.080)	(0.077)	(0.079)	(0.078)
N	5,383,850	5,383,850	5,383,850	12,545,295	12,545,295	12,545,295
R^2	0.004	0.004	0.004	0.006	0.006	0.006
$w_{ijt} = 0$	No	No	No	Yes	Yes	Yes
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B10: Risk Aversion and Bias - Summary Statistics

The table reports the summary statistics of the parameter estimates $\hat{\alpha}_i$ and $\hat{\gamma}_i$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The reported results are obtained under measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights.

	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.007	1.236
standard deviation	0.068	5.850
median	0.010	1.117
min	-0.676	-44.666
max	0.736	48.631
skewness	-0.626	1.075
kurtosis	27.395	13.200

B.2 Figures

Figure B1: Explained R^2

The figure reports the fraction of variation in expected excess returns explained by various fixed effects. For (1), (2) and (3) we report the R^2 of the following regression $\mu_{i,j,t} - r_f = H_k + \epsilon_{i,j,t}$. (1) reports results for manager fixed effects, i.e., $H_k = H_i$; (2) for stock fixed effects $H_k = H_j$; (3) for time fixed effects $H_k = H_t$. (4) reports the R^2 for separate manager, stock and time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_j + H_t + \epsilon_{i,j,t}$. (5) reports the results for manager-time and stock fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_j + \epsilon_{i,j,t}$. (6) reports the results for manager and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_{j,t} + \epsilon_{i,j,t}$. (7) reports the results for manager-time and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$.

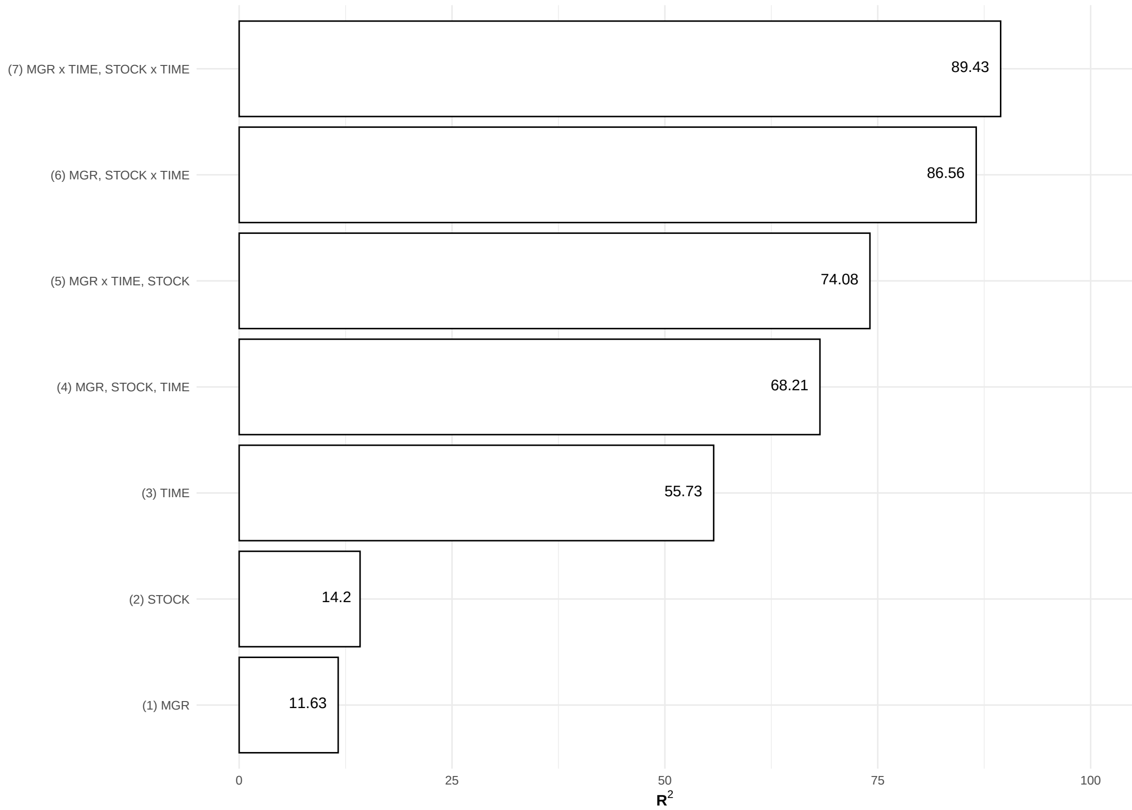


Figure B2: Managers' Careers

The upper panel shows the distribution of starting date for the managers' careers, as the first date we can track the manager in sample. The bottom panel shows the distribution of tenure across managers and dates as the difference between the current date and the starting date in quarters.

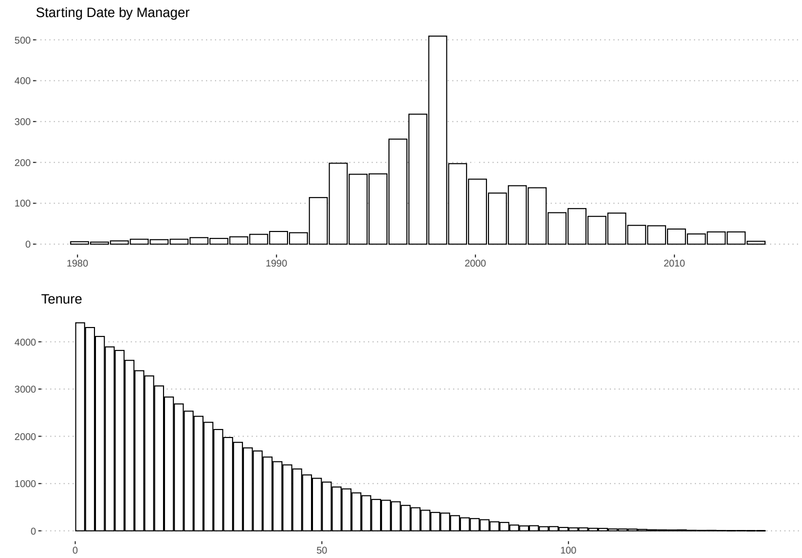


Figure B3: Stock-Manager Experience

The upper panel depicts the starting date of each manager-stock pair $t_{i,j,0}$, as the first date in which we observe a certain manager i holding a certain stock j. The middle panel shows the distribution of stock-manager experience, i.e., for any date t, manager i and stock j experience$_{i,j,t} = t - t_{i,j,0}$. The bottom panel reports the distribution of the maximal experience achieved for each manager-stock pair, i.e., max. experience$_{i,j} = \max_t\{\text{experience}_{i,j,t}\}$.

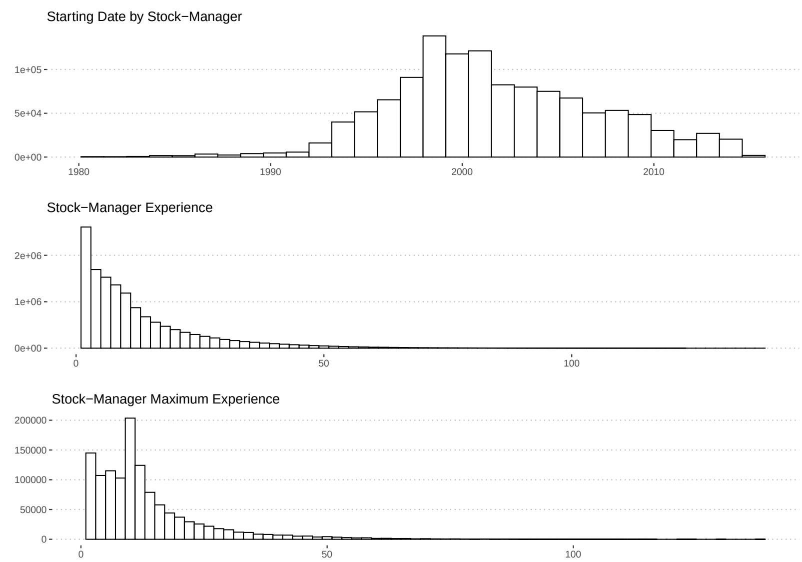


Figure B4: Weights on Past Experience

The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included in the upper panel, a manager-stock pair must have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

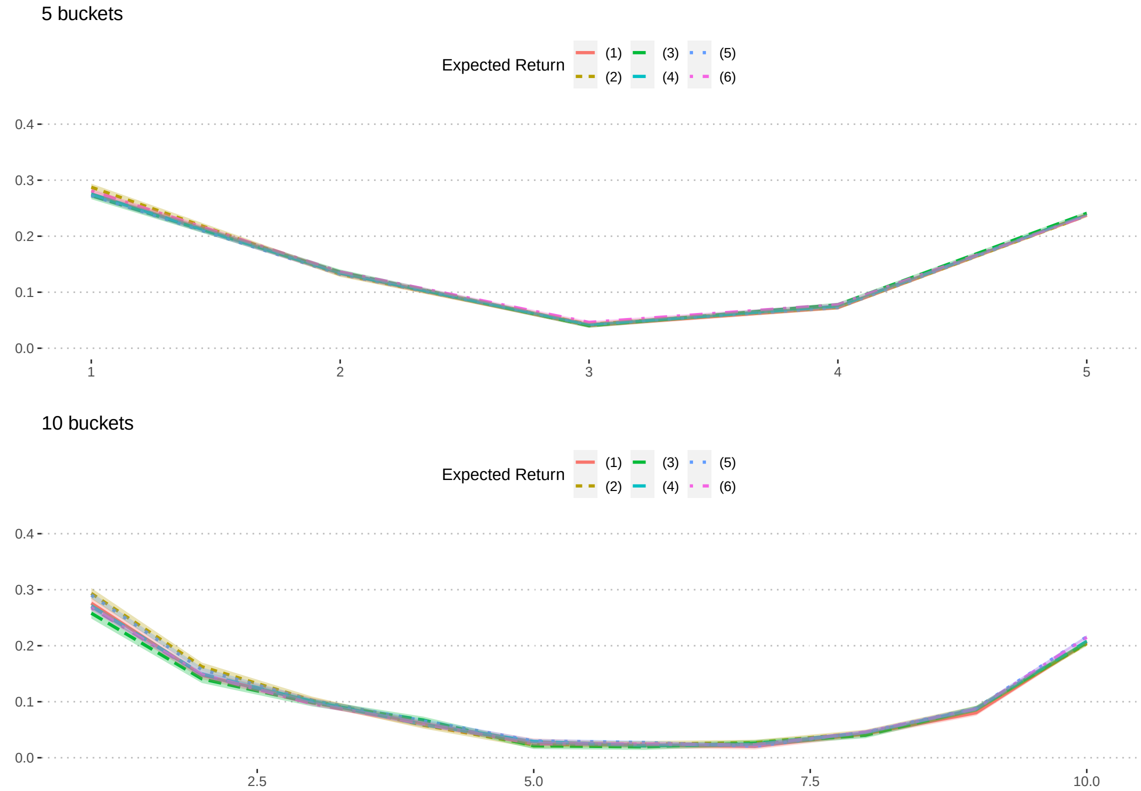


Figure B5: Weights on Past Experience by Number of Managers

The figure reports the parameter estimates for $\beta_{q,n}$ obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The horizontal axis refers to q, while each line to $n_{i,t} \in \{1, 2, 3, 4 \text{ or more}\}$. The top row reports the results for $Q = 5$, the bottom for $Q = 10$. The left column plots coefficients for measure (1), namely expected excess returns are computed without $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$; the right column for measure (4), namely expected excess returns are computed with $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$.

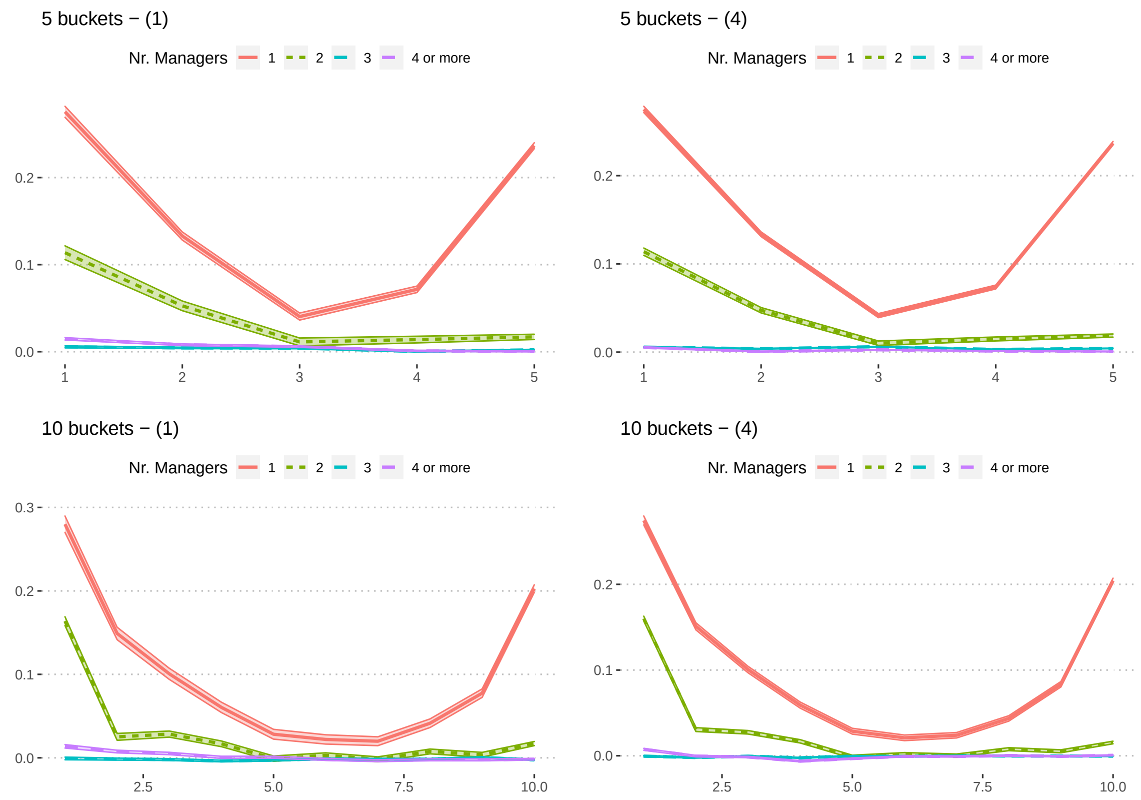


Figure B6: Weights on Past Experience - Managers Who Have Switched Funds

The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. In the upper panel, manager-stock pairs have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

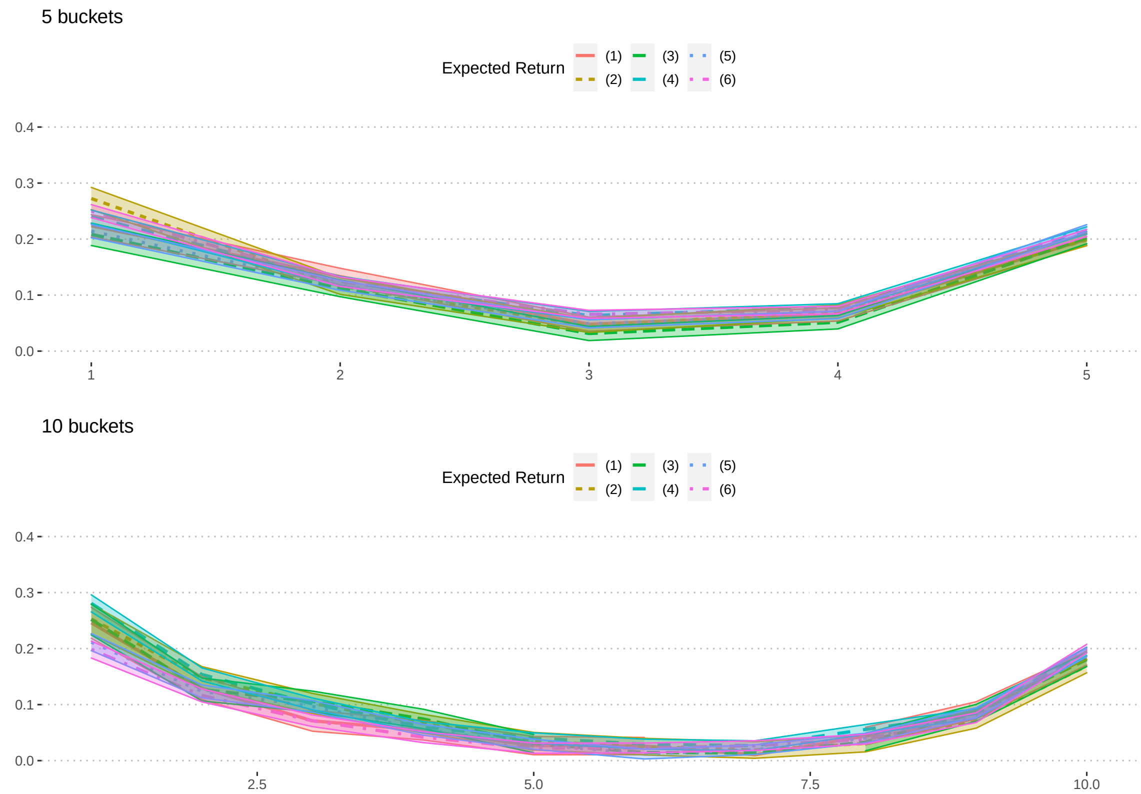


Figure B7: Weighting Functions - Various Examples

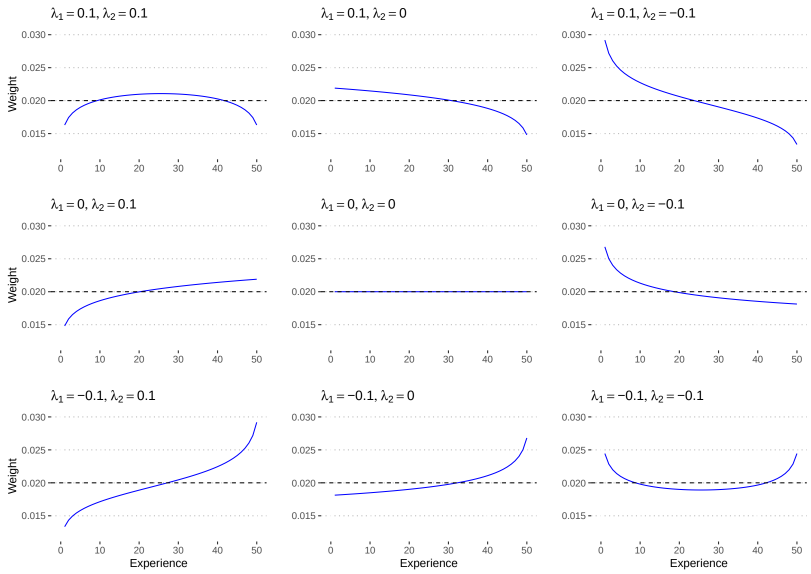


Figure B8: Empirical Weighting Function

The figure plots the weights implied by the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k} \right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t according to measure (1), $r_{j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. The upper panel reports weights for a manager with stock-specific experience of 9 quarters and the lower for 13 quarters.

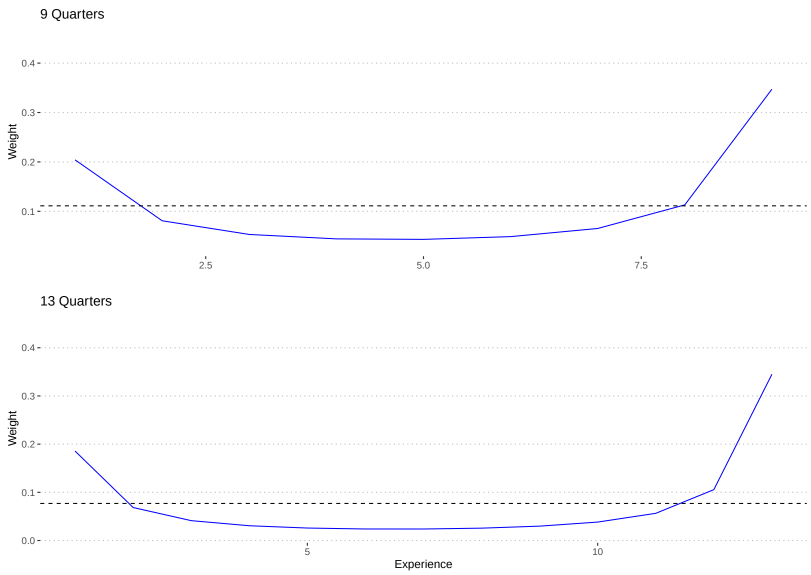


Figure B9: Bias and Risk Aversion

The figure shows the empirical distribution of the parameter estimates $\hat{\alpha}_{i,t}$ and $\hat{\gamma}_{i,t}$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The dashed lines represent the median bias and risk aversion, respectively. The histograms are trimmed for outliers.

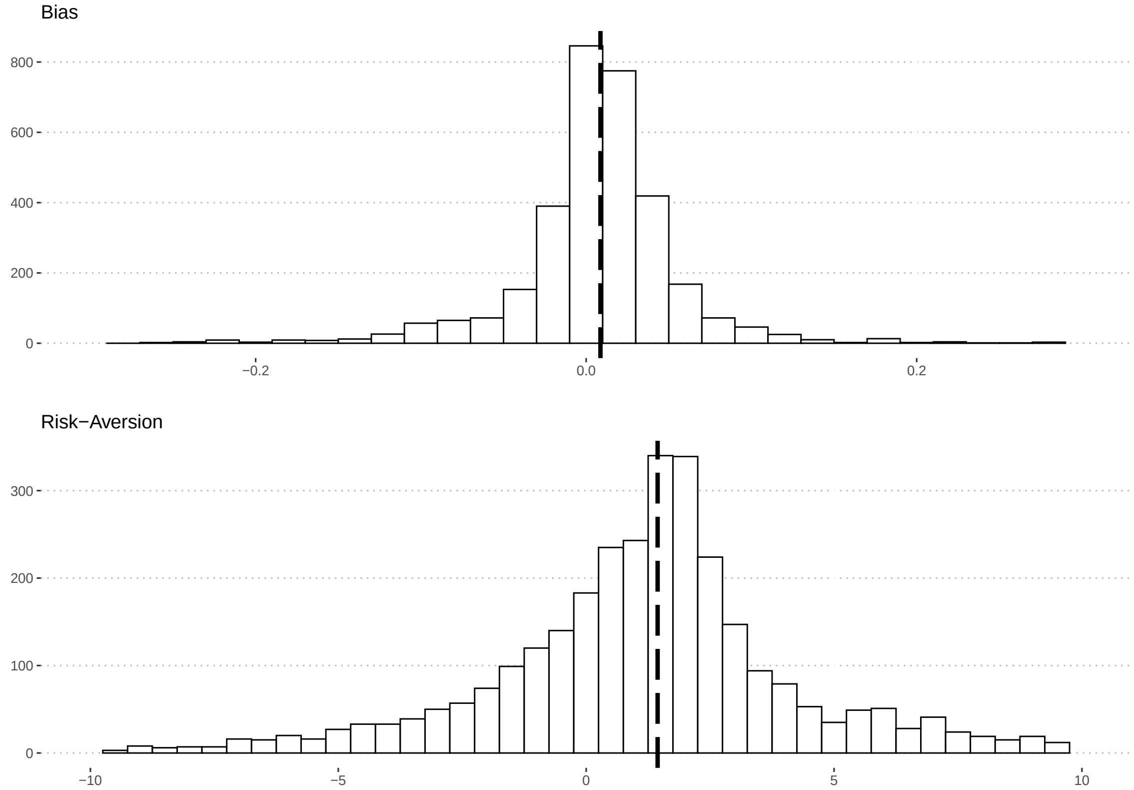


Figure B10: Bias and Risk Aversion by Tenure

The figure plots the parameter estimates $\hat{\alpha}_\tau$ and $\hat{\gamma}_\tau$ obtained by running one regression per tenure τ with the following specification: $r_{j,t+1} - r_f = \alpha_\tau + \gamma_\tau(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i},\boldsymbol{t}})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i},\mathbf{t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated parameter $\hat{\alpha}_\tau$, while Risk Aversion is the estimated parameter $\hat{\gamma}_\tau$. Tenure is measured in quarters since the first observation where we can identify the manager. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.

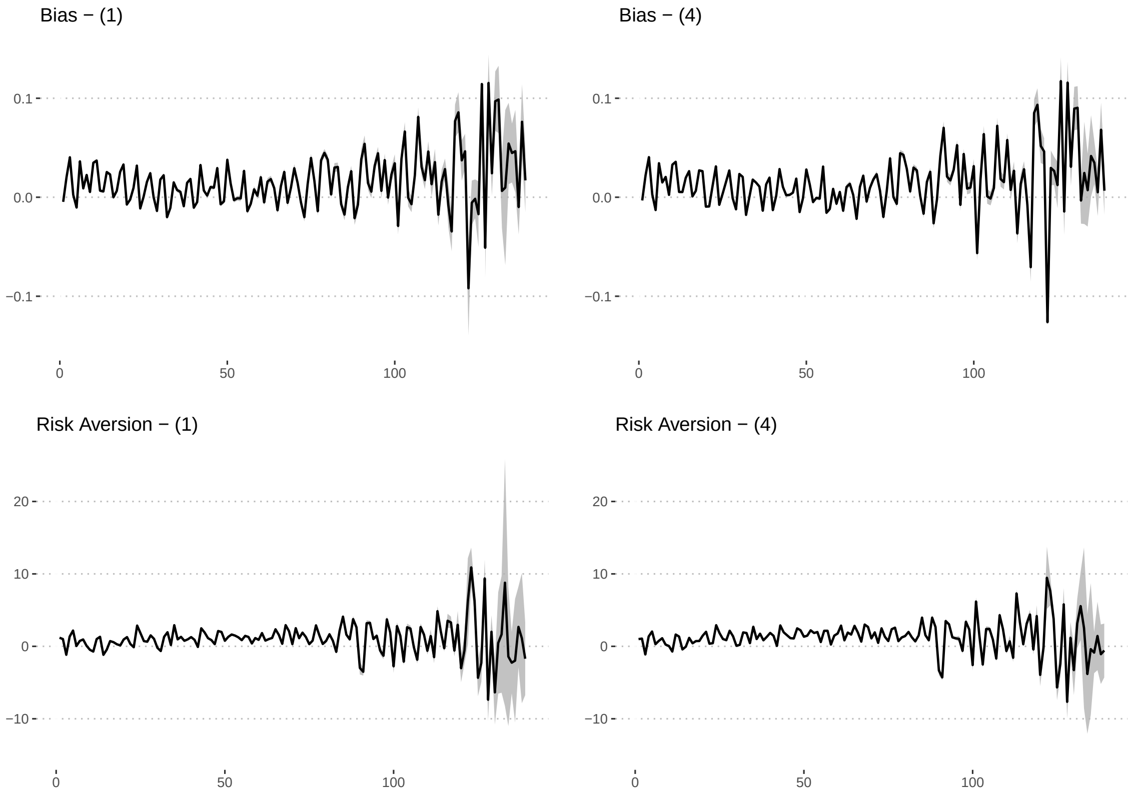
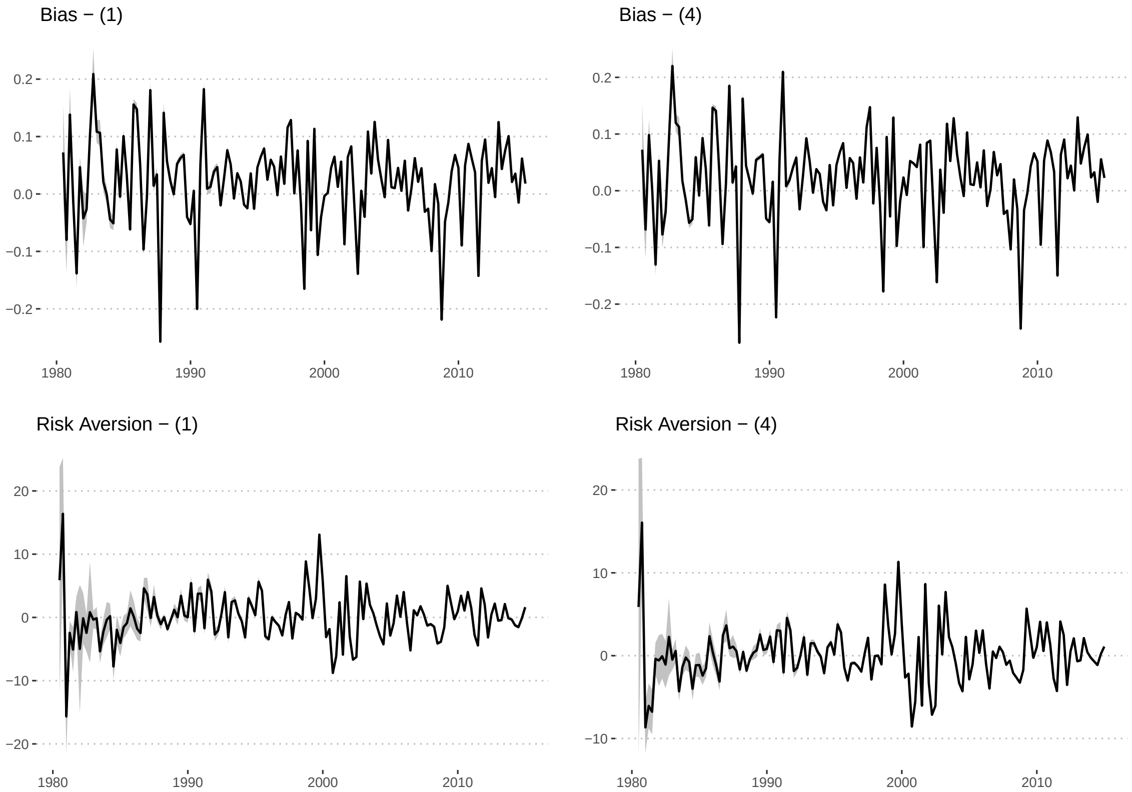


Figure B11: Bias and Risk Aversion by Date

The figure plots the parameter estimates $\hat{\alpha}_t$ and $\hat{\gamma}_t$ obtained by running one regression per date with the following specification: $r_{j,t+1} - r_f = \alpha_t + \gamma_t(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated paramater $\hat{\alpha}_t$, while Risk Aversion is the estimated parameter $\hat{\gamma}_t$. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.



B.3 Optimal Portfolio Choice

In what follows we provide four examples of optimal portfolio choice and describe how we can (or cannot) achieve identification of beliefs. We first look at an investor facing borrowing constraints, second an investor facing short sale constraints, third we look at an investor worried about model misspecification and, finally, an investor who is tracking a benchmark. We show that we can identify beliefs in the first three cases, while the last one requires us to make additional assumptions.

B.3.1 Borrowing Constraint

We follow the approach of Cvitanic and Karatzas (1992), Xu and Shreve (1992) and Tepla (2000). There exists a standard filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P})$ where all the *usual* regularity conditions are satisfied. We assume that the investor maximises his expected utility over terminal wealth $\mathbb{E}_0[U(W_T)]$. Returns follow a geometric Brownian motion and the investor faces a borrowing constraint. He solves the following problem:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \quad \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{B.1}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.2}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.3}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.4}$$

$$\boldsymbol{w_t'}\mathbf{1} \leq k \tag{B.5}$$

where B_t is the price of a risk-free bond, $\boldsymbol{S_t}$ is a vector of stock prices, $\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \left[\frac{dS_{1,t}}{S_{1,t}}, \dots, \frac{dS_{j,t}}{S_{j,t}}, \dots, \frac{dS_{N,t}}{S_{,N,t}}\right]'$, r_f is the instantaneous risk-free rate, $\boldsymbol{\mu_t}$ is the vector of stock return drifts, $\boldsymbol{w_t}$ is the vector of stock portfolio weights, $\Sigma_t^{\frac{1}{2}}$ is the matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, $\mathbf{1}$ is a vector of ones and k is a real number. Cvitanic and Karatzas (1992) show that the problem in (B.1)-(B.5) is equivalent to an unconstrained problem with modified drifts, i.e., where (B.2) and (B.3) are replaced by:

$$\frac{dB_t}{B_t} = (r_f + \delta(\boldsymbol{v_t}))dt \tag{B.6}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = (\boldsymbol{\mu_t} + \boldsymbol{v_t} + \delta(\boldsymbol{v_t})\mathbf{1})dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{B.7}$$

where the support function $\delta(\boldsymbol{x}) = \sup_{\boldsymbol{w}'\mathbf{1}\leq k}(-\boldsymbol{w}'\boldsymbol{x})$, $\boldsymbol{v_t}$ is such that $\delta(\boldsymbol{v_t}) < \infty$. Cvitanic and Karatzas (1992) show that the optimal $\boldsymbol{v_t^*}$ and portfolio weights $\boldsymbol{w_t^*}$ can be obtained by solving the 'dual' Hamilton-Jacobi-Bellman equation[1]. In particular, the optimal portfolio weights are:

$$\boldsymbol{w_t^*} = \frac{1}{\gamma}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1} - \boldsymbol{v_t^*}) \tag{B.8}$$

where $\boldsymbol{v_t^*} = \arg\min_{\{\boldsymbol{v}\text{ s.t. }\delta(\boldsymbol{v})<\infty\}}\left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}}\boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t})\right]$ and $\boldsymbol{\theta_t} = \Sigma_t^{-\frac{1}{2}}(\boldsymbol{\mu_t} - r_f\mathbf{1})$. Tepla (2000) shows that $\boldsymbol{v_t^*} = \bar{v}^*\mathbf{1}$ with $\bar{v}^* = \frac{\gamma(1-\gamma)-\mathbf{1}'\Sigma_t^{-1}(\boldsymbol{\mu_t}-r_f\mathbf{1})}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}}$ when the borrowing constraint binds, and zero otherwise. Notice that the above result implies that the solution to the constrained optimisation problem is equivalent to that of an unconstrained problem with a risk-free rate shifted by the scalar $\bar{v}^*$. Identification of beliefs is easily achieved in (B.8) by saturating the model with manager-time fixed effects in order to absorb any variation in manager-specific borrowing constraints. Specifically, for each manager i solving the above problem, the subjective beliefs can be expressed as:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma_i\Sigma_t\boldsymbol{w_{i,t}^*} + \boldsymbol{H_{i,t}} \tag{B.9}$$

where the manager-time fixed effect is equal to $\boldsymbol{H_{i,t}} = \bar{v}_i{}^*\mathbf{1}$.

[1]See Sections 12 and 15 of Cvitanic and Karatzas (1992). In particular, see equations (15.1), (15.2) and (15.10).

B.3.2 Short Sale Constraints

The manager solves the following problem[2]:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{B.10}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.11}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.12}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.13}$$

$$-w_{j,t} \leq 0 \quad \forall j = 1, 2, ..., N \tag{B.14}$$

The problem (B.10)-(B.14) can be solved by using Cvitanic and Karatzas (1992) and Xu and Shreve (1992)'s dual approach, similarly to the previous section. The support function now becomes $\delta(\boldsymbol{x}) = \sup_{\{-w_{j,t}\leq 0 \quad \forall j=1,2,...,N\}}(-\boldsymbol{w}'\boldsymbol{x})$. As before, we can find $\boldsymbol{v_t^*}$ by solving:

$$\min\left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}}\boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t})\right] \text{ s.t.} \tag{B.15}$$

$$-\boldsymbol{v_t} \leq \mathbf{0} \tag{B.16}$$

Denote the vector of Lagrange multipliers on the the constraint in equation (B.16) by $\boldsymbol{\lambda_t} = [\lambda_{1,t}, ..., \lambda_{N,t}]'$. Taking first-order conditions of the above minimisation problem yields:

$$\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1} + \boldsymbol{v_t^*}) + \boldsymbol{\lambda_t} = \mathbf{0} \tag{B.17}$$

Consider the following partitions: $\boldsymbol{v_t^*} = [\mathbf{0}' \quad \boldsymbol{v_t^{(2)*'}}]'$, $\boldsymbol{\lambda_t} = [\boldsymbol{\lambda_t^{(1)'}} \quad \mathbf{0}']'$, where we have divided between assets for which the short sale constraint does not bind and those for which it does. We can also partition the vector of expected excess returns

[2]This problem is similar to the discrete problem analyzed by Koijen and Yogo (2019) as $\gamma \to 1$.

and the covariance matrix: $\boldsymbol{\mu_t} - r_f\mathbf{1} = [(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})' \quad (\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1})']'$,

$$\Sigma_t = \begin{bmatrix} \Sigma_t^{(1,1)} & \Sigma_t^{(1,2)} \\ \Sigma_t^{(2,1)} & \Sigma_t^{(2,2)} \end{bmatrix},$$

Standard results imply that the inverse of the covariance matrix can be partitioned as:

$$\Sigma_t^{-1} = \begin{bmatrix} \Omega_t^{(1)} & -\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)} \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)} & \Omega_t^{(2)} \end{bmatrix}$$

where

$$\Omega_t^{(1)} = \left(\Sigma_t^{(1,1)} - \Sigma_t^{(1,2)}\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\right)^{-1}$$

$$\Omega_t^{(2)} = \left(\Sigma_t^{(2,2)} - \Sigma_t^{(2,1)}\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\right)^{-1}$$

Using the above, rewrite equation (B.17) as:

$$\mathbf{0} = \begin{bmatrix} \Omega_t^{(1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) + \boldsymbol{\lambda_t^{(1)}} \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) + \Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) \end{bmatrix} \quad \text{(B.18)}$$

Multiplying the second row of (B.18) by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding it to the first row allows us to solve for the Lagrange multipliers:

$$\boldsymbol{\lambda_t^{(1)}} = -\Sigma_t^{(1,1)-1}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) \quad \text{(B.19)}$$

Insert the multipliers into the first-order condition in equation (B.17) to obtain:

$$\boldsymbol{v_t^*} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{v_t^{(2)*}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) - \left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1}\right) \end{bmatrix} \quad \text{(B.20)}$$

We can now substitute $\boldsymbol{v_t^*}$ into equation (B.8) and solve for the optimal weights:

$$\boldsymbol{w_t^*} = \begin{bmatrix} \boldsymbol{w_t^{(1)*}} \\ \mathbf{0} \end{bmatrix} = \frac{1}{\gamma}\begin{bmatrix} \Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)}\left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})\right) \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) + \Omega_t^{(2)}\left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})\right) \end{bmatrix} \quad \text{(B.21)}$$

Multiplying the second row by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding the two rows together gives the optimal weights on the unconstrained assets:

$$\boldsymbol{w}_t^{(1)*} = \frac{1}{\gamma}\Sigma_t^{(1,1)-1}\left(\boldsymbol{\mu}_t^{(1)} - r_f\mathbf{1}\right) \tag{B.22}$$

Intuitively, the optimisation program of a short sale constrained investor results in an unconstrained portfolio allocation over the set of assets for which the constraint does not bind. For each manager i, identification of beliefs can be achieved by inverting equation (B.22):

$$\boldsymbol{\mu}_{i,t}^{(1)} - r_f\mathbf{1} = \gamma_i\Sigma_t^{(1,1)}\boldsymbol{w}_{i,t}^{(1)*} \tag{B.23}$$

B.3.3 Model Misspecification

In this section we follow the approach of Maenhout (2004) and analyse the behaviour of an investor worried about model misspecification. The investor solves the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}, C_s\}} \mathbb{E}_0\left[\int_0^\infty f(C_s, J_s)ds\right] \quad \text{s.t.} \tag{B.24}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.25}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t}dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{B.26}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t}'\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) - \frac{C_t}{W_t}dt \tag{B.27}$$

where we are vague about the functional form of the value function. Standard dynamic optimisation arguments yield the following HJB equation:

$$0 = \sup_{\{\boldsymbol{w_t}, C_t\}} \{f(C_t, J_t)dt + \mathbb{E}_t\left[dJ_t\right]\} \tag{B.28}$$

Equation (B.28) assumes that the investor is certain about the value of $\mathbb{E}_t\left[dJ_t\right]$ and chooses his portfolio accordingly. An investor worried about model misspecification will choose the optimal allocation given the worst-case scenario. Following Anderson et al. (2003), Maenhout (2004) shows that the wealth of the investor under the endogenously chosen model for $u(W_t)$ will evolve according to:

$$dW_t = W_t \left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f\mathbf{1}) - \frac{C_t}{W_t} \right) dt + W_t \boldsymbol{w_t'} \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} + W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} u(W_t) dt \tag{B.29}$$

where $u(W_t)$ is a drift term chosen by the investor to minimise the following expression:

$$u^*(W_t) = \inf_{u_t} \left\{ \mathbb{E}_t[dJ_t|u_t] + \frac{1}{2\Psi} u_t^2 W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} dt \right\} \tag{B.30}$$

where $\mathbb{E}_t[dJ_t|u_t]$ is computed under the law of motion in equation (B.29). Among all the models for $u(W_t)$ the investor chooses the least favourable one in terms of its effect on $\mathbb{E}_t[dJ_t|u_t]$, subject to the entropy constraint $\frac{1}{2\Psi} u_t^2 W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} dt$. The HJB equation thus becomes:

$$\begin{aligned} 0 = \sup_{\{\boldsymbol{w_t}, C_t\}} \inf_{u_t} \quad & f(C_t, J_t) + \frac{\partial J_t}{\partial t} + J_{W_t} W_t \left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f\mathbf{1}) - \frac{C_t}{W_t} \right) + \\ & J_{W_t} W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} u_t + \frac{1}{2\Psi} u_t^2 W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} + \frac{1}{2} J_{W_t W_t} W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} \end{aligned} \tag{B.31}$$

The agent will choose $u(W_t)^* = -J_{W_t}\Psi$. The optimal portfolio, therefore, will be:

$$\boldsymbol{w_t^*} = -\frac{J_{W_t}}{[J_{W_t W_t} - J_{W_t}^2 \Psi] W_t} \Sigma_t^{-1} (\boldsymbol{\mu_t} - r_f\mathbf{1}) \tag{B.32}$$

An investor concerned about model misspecification will behave like an otherwise identical investor with relative risk aversion of $\gamma_{i,t} = -\frac{[J_{W_{i,t}W_{i,t}} - J_{W_{i,t}}^2 \Psi_i] W_{i,t}}{J_{W_{i,t}}}$. In this case, identification follows in a way similar to the standard model presented in the main text.

B.3.4 Benchmarking

In the spirit of van Binsbergen et al. (2008), consider an investor who has his objective function defined over his terminal wealth W_T relative to a benchmark portfolio

M_T. He will solve the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}\}} \mathbb{E}_0 \left[f\left(\frac{W_T}{M_T^{\beta}}\right)\right] \quad \text{s.t.} \tag{B.33}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.34}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.35}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.36}$$

Assume that the benchmark has weights $\boldsymbol{\theta_t}$ in the N risky assets and therefore evolves according to:

$$\frac{dM_t}{M_t} = \frac{dB_t}{B_t} + \boldsymbol{\theta_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.37}$$

The problem can be recast in terms of the state variable $X_t = \frac{W_t}{M_t^{\beta}}$ with the following law of motion:

$$\begin{aligned}\frac{dX_t}{X_t} =& ((1-\beta)r_f + (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'(\boldsymbol{\mu_t} - r_f\mathbf{1}))dt - \frac{1}{2}\beta(\beta-1)\boldsymbol{\theta_t'}\Sigma_t\boldsymbol{\theta_t}dt + \\ & (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} - (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t\beta\boldsymbol{\theta_t}dt\end{aligned} \tag{B.38}$$

If we set up the HJB equation and take first-order conditions, we obtain the optimal weights:

$$\boldsymbol{w_t^*} = -\frac{J_{X_t}}{J_{X_tX_t}X_t}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1}) + \beta\boldsymbol{\theta_t}\left(1 + \frac{J_{X_t}}{J_{X_tX_t}X_t}\right) \tag{B.39}$$

In this case, it is not obvious that we can identify beliefs. However, if there is no variation in the objective function in the cross-section of managers adopting the same benchmark portfolio $\boldsymbol{\theta_t}$, stock-time fixed effects would suffice to recover expectations. Although the above model requires an additional assumption to achieve identification, this is consistent with the common practice of evaluating managers using summary statistics such as CAPM alphas (Berk and van Binsbergen, 2016; Barber et al., 2016). For instance, set $f\left(\frac{W_T}{M_T^{\beta}}\right) = \frac{1}{1-\gamma}\left(\frac{W_T/W_0}{(M_T/M_0)^{\beta}}\right)^{1-\gamma} =$

$\frac{1}{1-\gamma}\left(\frac{R_{W,T}}{R_{M,T}^{\beta}}\right)^{1-\gamma}$. That would be equivalent to solving:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}] - \frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T}) \tag{B.40}$$

where $r_{W,T} = \log W_T/W_0$ and $r_{M,T} = \log M_T/M_0$ are log-returns. The manager is maximising $\alpha = \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}]$ subject to the tracking error penalisation $\frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$[3]. In this case $-\frac{J_{X_t}}{J_{X_tX_t}X_t} = 1/\gamma$ and we could recover beliefs using:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma\Sigma_t\boldsymbol{w_{i,t}^*} + \boldsymbol{H_t} \tag{B.41}$$

Notice that each element of the vector $\boldsymbol{H_t}$ varies at the stock-time level, i.e.: $\boldsymbol{H_t} = (1-\gamma)\beta\Sigma_t\boldsymbol{\theta_t}$.

[3]As it is well known, the agent penalises tracking error for any value of $\gamma > 0$, even for $0 < \gamma \leq 1$. To see this, notice that we can substitute $\mathbb{E}_0[r_{W,T} - \beta r_{M,T}] = \log\mathbb{E}_0\left[\frac{R_{W,T}}{R_{M,T}^{\beta}}\right] - \frac{1}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$ and obtain the following objective:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \log\mathbb{E}_0\left[\frac{R_{W,T}}{R_{M,T}^{\beta}}\right] - \frac{\gamma}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$$

B.4 Data Construction

In this section we provide details on the construction of the data that are used in the paper. We start with the universe of mutual funds in the CRSP database. We remove funds whose manager name clearly does not refer to a person[4]. After having obtained a list of names of managers, we look for cases in which the same manager is spelled differently, e.g. "John Smith", "J. Smith", "J Smith" or just "Smith". To be sure that the pairing is done correctly we proceed in the following way: first, we compute a matrix of distances between names using cosine, Jaccard and Jaro-Winkler methods. We then keep pairs that have a distance below a distance-specific threshold (0.10, 0.17, 0.10 for the cosine, Jaccard and Jaro-Wrinkler methods, respectively) that is set to make sure that we avoid false negatives. We then proceed to manually check over 15,000 pairs to guarantee proper matching with the help of online resources and common sense. After having obtained a list of managers with the dates in which they manage a specific fund, we follow Evans (2010) and Benos et al. (2010) to screen for equity mutual funds. First, if available, we keep funds with the following Lipper class: EIEI, G, LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, SCVE. We then keep the funds with missing Lipper class and the following Strategic Insight Objective Code: AGG, GMC, GRI, GRO, ING, SCG. If neither of the previous are available, we use the following Wiesenberger Fund Type Codes: G, G-I, AGG, GCI, GRI, GRO, LTG, MCG, and SCG. We then keep all the funds with policy equal to CS. Finally, we remove funds with less than 80% of holdings in common equity, similarly to Kacperczyk et al. (2006). To check for possible mistakes we keep funds with CRSP objective code starting with E and M and remove those starting with EF. This provides us with a manager-by-manager history of the funds managed that we subsequently match with the S12 type1 file from the Thomson-Reuters Institutional Holdings database, using Russ Wermer's MFLinks tables. We then proceed by joining with the S12 type2 and type3 files to obtain a history of holdings.

[4]We use various automatic screens like "advisors", "ltd", "limited", etc..., paired with manual inspection.

We continue by adding stock return and balance sheet data using CRSP and Compustat, respectively. From the CRSP Compustat Merged Database we select LinkTypes LU and LC and LinkPrim P and C for stocks with share codes of 10 and 11. After we have merged the two datasets, we compute dividends using CRSP returns and returns not including distributions, similarly to Koijen and Yogo (2019). From Compustat we compute the following quantities: *me* as market equity, *beme* as the book to market equity ratio, *dp* as the ratio between dividends and market prices, *profitability* as the ratio between operating profits and book equity and *investment* as the growth rate of assets similarly to Fama and French (2015).

We then proceed with the construction of the scaled demands $\hat{\Sigma}_t \boldsymbol{w_{i,t}}$. We start from CRSP daily return data and compute covariance matrices using the previous year. We compute three daily covariance matrices: $\hat{\Sigma}_t^{d,1}$ which is the sample covariance matrix, and two Bayesian shrinkage estimates. The first one follows Touloumis (2015) and shrinks the daily sample covariance towards a target diagonal matrix with the sample variances on the diagonal, i.e., the resulting estimator is $\hat{\Sigma}_t^{d,2} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_t^{target}$, with $\Sigma_t^{target} = I_N * \hat{\Sigma}_t^{d,1}$, where $*$ denotes the Hadamard product and I_N is an $N \times N$ identity matrix with N being the number of stocks. The third covariance estimator follows Ledoit and Wolf (2004) and shrinks the daily covariance matrix towards a diagonal matrix with the average variance on the diagonal, i.e., $\hat{\Sigma}_t^{d,3} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_t^{target}$, where $\tilde{\Sigma}_t^{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the daily sample covariance matrix. The shrinkage intensity λ is chosen similarly to Touloumis (2015) to minimise the risk function $\mathbb{E}[||\hat{\Sigma}_t^{d,k} - \Sigma_t^d||_F^2]$ where $||S||_F^2 = \frac{tr(S'S)}{dim(S)}$ denotes the Frobenius norm of matrix S, which results in $\lambda = \frac{Y_{2,T} + Y_{1,T}^2}{TY_{2,T} + \frac{N-T+1}{N} Y_{1,T}^2}$, where $Y_{1,T} = \frac{1}{T}\sum_{s=1}^T X_s' X_s - \frac{1}{P_2^T}\sum_{s \neq h} X_h' X_s$, $Y_{2,T} = \frac{1}{P_2^T}\sum_{s \neq h}(X_h' X_s)^2 - 2\frac{1}{P_3^T}\sum_{s \neq h \neq k} X_s' X_h X_s' X_k + \frac{1}{P_4^T}\sum_{s \neq h \neq k \neq w} X_s X_h' X_k X_w'$ with X_j being the vector of stock returns for which we have T observations and $P_a^b = \frac{b!}{(b-a)!}$. Finally, we can scale the matrices $\hat{\Sigma}_t^{d,k}$ by the average number of trading days in a quarter, which in our sample is equal to $\frac{num.obs}{num.quarters} = 63.07$ to obtain our quarterly estimators $\hat{\Sigma}_t^k = \frac{num.obs}{num.quarters} \times \hat{\Sigma}_t^{d,k}$. We can then proceed to compute scaled demands as $\hat{\Sigma}_t^k \boldsymbol{w_{i,t}}$. We compute two vectors of scaled demands for each estimator: one that does not include stocks that currently have zero weights, but belong to the investment opportunity set of the manager, and one that does, i.e., in the first case all the $w_{i,j,t}$ in $\boldsymbol{w_{i,t}}$ are different from zero, while in the second $\boldsymbol{w_{i,t}}$ has some zero elements. The investment opportunity set is constructed similarly to Koijen and Yogo (2019) and includes all stocks that are currently held or have ever been held by the manager in the past 11 quarters.

B.5 Parametric Estimation

As described in Section 2.5, we estimate the model in equation (2.22) via non-linear least squares (NLS). In particular we obtain the coefficients $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)'$ by minimising the sum of squared errors:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_i \sum_j \sum_t \left(\mu_{i,j,t} - r_f - \beta \left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}} r_{j,t+1-k} \right) - H_{i,t} - H_{j,t} \right)^2 \tag{B.42}$$

We perform the minimisation with $(\hat{\lambda}_1, \hat{\lambda}_2) \in [-5, 5] \times [-5, 5]$ via Simulated Annealing and limited-memory BFGS[5]. Fixed effects are partialled out by demeaning $\mu_{i,j,t} - r_f$ and $\left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}} r_{j,t+1-k} \right)$. To compute standard errors, we can rewrite (B.42) as:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \sum_{p=1}^{P} (y_p - \varphi(\boldsymbol{x_p}; \boldsymbol{\theta}))^2 \tag{B.43}$$

where the index p is a short-hand for all the P combinations of i, j, t. We next follow the approach of Davidson and MacKinnon (2001) and recover standard errors using Gauss-Newton Regressions. Consider the 3×1 gradient vector $\Psi(\boldsymbol{x_p}; \boldsymbol{\theta}) = \frac{\partial \varphi(\boldsymbol{x_p}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and the following regression:

$$y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}}) = \Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})' \boldsymbol{b} + u_p \tag{B.44}$$

where we regress the residuals $y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$ on the estimated gradient $\Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$[6]. Denote the $P \times 3$ matrix of gradient observations as $\hat{\Psi} = [\Psi(\boldsymbol{x_1}; \hat{\boldsymbol{\theta}}), ..., \Psi(\boldsymbol{x_P}; \hat{\boldsymbol{\theta}})]'$, then we can estimate the covariance matrix of the coefficients $\boldsymbol{b}$ using the standard clustered "sandwich" estimator:

$$S(\hat{\boldsymbol{b}}) = (\hat{\Psi}'\hat{\Psi})^{-1} \hat{\Psi}' \hat{\Omega} \hat{\Psi} (\hat{\Psi}'\hat{\Psi})^{-1} \tag{B.45}$$

[5]Notice that, conditional on λ_1 and λ_2, β can be estimated via OLS and, therefore, is left unconstrained.

[6]For expositional reasons we exclude the estimated fixed effects from $\boldsymbol{\theta}$. Given that they enter linearly in $\varphi(\boldsymbol{x_p}; \boldsymbol{\theta})$, their gradients are identical to the matrix containing the full set of dummies and, therefore, can be taken care of by including dummies on the right hand side of (B.44) or by demeaning.

Davidson and MacKinnon (2001) show that the covariance matrix of $\boldsymbol{b}$ in (B.45) is a consistent estimator for the covariance of $\boldsymbol{\theta}$.

B.6 Additional Tables and Figures

Table B11: The Effect of Average Experienced Returns

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.148***	0.140***	0.146***	0.188***	0.179***	0.189***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
profitability	-0.002	-0.0010	-0.002	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.040***	0.032***	0.035***	0.051***	0.039***	0.042***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
BE/ME	0.012	0.020***	0.012*	0.016*	0.019**	0.017**
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
ME	0.011	0.009	0.012	0.009	0.0009	0.008
	(0.015)	(0.012)	(0.013)	(0.018)	(0.016)	(0.017)
D/P	-0.019***	-0.017***	-0.018***	-0.005	-0.006	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
N	1,153,333	1,153,333	1,153,333	2,596,853	2,596,853	2,596,853
R^2	0.591	0.583	0.588	0.546	0.538	0.536
Within-R^2	0.016	0.014	0.015	0.021	0.019	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B12: The Effect of Experienced Returns - Five Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.297***	0.283***	0.286***	0.274***	0.264***	0.277***
	(0.009)	(0.010)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.137***	0.129***	0.138***	0.125***	0.115***	0.121***
	(0.009)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)
β_3	0.061***	0.054***	0.057***	0.055***	0.048***	0.054***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.004)	(0.005)
β_4	0.084***	0.083***	0.088***	0.085***	0.078***	0.084***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_5	0.266***	0.259***	0.262***	0.267***	0.258***	0.261***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005*	0.0009	-0.005	-0.010**	-0.006	-0.008*
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
investment	0.006	0.003	0.002	0.019***	0.011*	0.013**
	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.066***	0.072***	0.062***	0.053***	0.056***	0.054***
	(0.014)	(0.015)	(0.016)	(0.010)	(0.010)	(0.010)
ME	-0.009	-0.007	-0.005	-0.012	-0.017	-0.013
	(0.015)	(0.012)	(0.013)	(0.020)	(0.019)	(0.019)
D/P	-0.008	-0.003	-0.006	0.005	0.003	0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
N	724,999	724,999	724,999	1,783,648	1,783,648	1,783,648
R^2	0.594	0.587	0.591	0.556	0.547	0.545
Within-R^2	0.066	0.064	0.065	0.070	0.067	0.069
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B13: The Effect of Experienced Returns - Ten Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.259***	0.224***	0.237***	0.235***	0.229***	0.242***
	(0.011)	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)
β_2	0.123***	0.116***	0.131***	0.124***	0.115***	0.120***
	(0.007)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.098***	0.094***	0.102***	0.088***	0.083***	0.084***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)
β_4	0.078***	0.064***	0.073***	0.069***	0.063***	0.065***
	(0.006)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
β_5	0.059***	0.048***	0.049***	0.047***	0.041***	0.043***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_6	0.061***	0.053***	0.055***	0.053***	0.053***	0.052***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_7	0.067***	0.066***	0.066***	0.066***	0.057***	0.063***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)
β_8	0.074***	0.063***	0.067***	0.071***	0.070***	0.074***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_9	0.107***	0.109***	0.113***	0.120***	0.114***	0.112***
	(0.005)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_{10}	0.243***	0.239***	0.239***	0.243***	0.247***	0.246***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.009	-0.011*
	(0.004)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
investment	-0.015*	-0.011	-0.015*	-0.005	-0.011*	-0.007
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.076***	0.078***	0.065***	0.069***	0.071***	0.066***
	(0.018)	(0.019)	(0.024)	(0.014)	(0.013)	(0.013)
ME	-0.019	-0.014	-0.011	-0.022	-0.028	-0.021
	(0.016)	(0.015)	(0.015)	(0.023)	(0.021)	(0.022)
D/P	-0.001	-0.003	-0.006	0.008	0.006	0.010
	(0.010)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)
N	403,968	403,968	403,968	980,175	980,175	980,175
R^2	0.598	0.588	0.596	0.567	0.557	0.555
Within-R^2	0.065	0.061	0.063	0.070	0.070	0.071
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B14: The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.283***	0.294***	0.288***	0.284***	0.288***	0.288***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)
β_2	0.077***	0.082***	0.078***	0.078***	0.079***	0.080***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
β_3	0.229***	0.231***	0.232***	0.231***	0.231***	0.233***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)
N	1,031,564	1,031,564	1,031,564	2,483,275	2,483,275	2,483,275
R^2	0.777	0.762	0.769	0.704	0.688	0.690
Within-R^2	0.039	0.041	0.040	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B15: The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.280***	0.273***	0.277***	0.273***	0.267***	0.277***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_2	0.073***	0.069***	0.071***	0.066***	0.060***	0.066***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)	(0.005)
β_3	0.236***	0.230***	0.233***	0.238***	0.229***	0.235***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.001	-0.000	-0.002*	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.019***	0.013*	0.014**	0.032***	0.021***	0.024***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.048***	0.056***	0.048***	0.043***	0.044***	0.044***
	(0.010)	(0.011)	(0.011)	(0.009)	(0.009)	(0.008)
ME	-0.002	-0.003	0.000	-0.006	-0.014	-0.007
	(0.014)	(0.012)	(0.012)	(0.019)	(0.017)	(0.017)
D/P	-0.009	-0.007	-0.008	0.004	0.002	0.003
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
N	937,382	937,382	937,382	2,258,925	2,258,925	2,258,925
R^2	0.582	0.573	0.578	0.545	0.536	0.535
Within-R^2	0.056	0.055	0.056	0.058	0.056	0.059
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B16: The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.016***	0.015***	0.022***	0.008***	0.008***	0.014***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
β_3	0.161***	0.159***	0.160***	0.166***	0.168***	0.165***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
N	618,451	618,451	618,451	1,499,594	1,499,594	1,499,594
R^2	0.812	0.799	0.807	0.744	0.729	0.733
Within-R^2	0.021	0.021	0.021	0.021	0.021	0.020
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B17: The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.208***	0.189***	0.198***	0.206***	0.194***	0.205***
	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
β_2	0.093***	0.083***	0.089***	0.077***	0.070***	0.076***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.215***	0.208***	0.209***	0.229***	0.225***	0.223***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.010	-0.011*
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
investment	-0.002	-0.003	-0.006	0.004	-0.001	0.001
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.056***	0.059***	0.048***	0.052***	0.051***	0.049***
	(0.013)	(0.015)	(0.018)	(0.012)	(0.011)	(0.011)
ME	-0.006	-0.002	0.001	-0.009	-0.012	-0.007
	(0.016)	(0.014)	(0.015)	(0.022)	(0.020)	(0.021)
D/P	-0.008	-0.009	-0.012	0.004	0.002	0.004
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
N	564,287	564,287	564,287	1,367,732	1,367,732	1,367,732
R^2	0.598	0.590	0.597	0.570	0.560	0.558
Within-R^2	0.042	0.039	0.040	0.046	0.044	0.045
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B18: The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.020***	0.017***	0.026***	0.014***	0.020***	0.019***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)
β_3	0.137***	0.131***	0.135***	0.136***	0.144***	0.141***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
N	343,058	343,058	343,058	753,526	753,526	753,526
R^2	0.870	0.864	0.866	0.834	0.821	0.824
Within-R^2	0.021	0.020	0.021	0.020	0.022	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B19: The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.168***	0.149***	0.157***	0.165***	0.152***	0.160***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)
β_2	0.067***	0.058***	0.065***	0.066***	0.063***	0.065***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
β_3	0.179***	0.173***	0.178***	0.183***	0.187***	0.186***
	(0.006)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)
profitability	-0.003	-0.003	-0.007	-0.016*	-0.014	-0.014
	(0.004)	(0.005)	(0.005)	(0.008)	(0.009)	(0.009)
investment	-0.029***	-0.027***	-0.029***	-0.029***	-0.035***	-0.032***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
BE/ME	0.093***	0.100***	0.092***	0.077***	0.074***	0.069***
	(0.027)	(0.027)	(0.026)	(0.019)	(0.019)	(0.018)
ME	-0.015	-0.007	-0.008	-0.023	-0.027	-0.021
	(0.019)	(0.017)	(0.018)	(0.027)	(0.025)	(0.026)
D/P	-0.001	0.004	-0.007	0.014	0.009	0.015
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)
N	314,557	314,557	314,557	691,634	691,634	691,634
R^2	0.671	0.661	0.666	0.655	0.644	0.644
Within-R^2	0.034	0.031	0.033	0.036	0.036	0.036
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B20: The Effect of Experienced Returns by Number of Managers

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.280***	0.164***	-0.001	0.014***	0.275***	0.160***	-0.001	0.008***
	(0.012)	(0.010)	(0.003)	(0.004)	(0.010)	(0.008)	(0.002)	(0.003)
β_2	0.149***	0.025***	-0.002	0.008**	0.151***	0.031***	-0.002	-0.000
	(0.008)	(0.007)	(0.002)	(0.004)	(0.006)	(0.004)	(0.002)	(0.002)
β_3	0.101***	0.028***	-0.002	0.005	0.101***	0.027***	-0.001	-0.001
	(0.006)	(0.004)	(0.002)	(0.003)	(0.004)	(0.003)	(0.001)	(0.002)
β_4	0.060***	0.017***	-0.004**	0.001	0.059***	0.017***	-0.002**	-0.006***
	(0.006)	(0.003)	(0.002)	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)
β_5	0.028***	-0.001	-0.003**	0.001	0.029***	-0.001	-0.000	-0.003**
	(0.005)	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)
β_6	0.022***	0.003	-0.001	-0.002	0.021***	0.002	-0.001	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
β_7	0.020***	-0.002	-0.002*	-0.003*	0.024***	0.000	0.000	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_8	0.041***	0.008***	-0.002	-0.002	0.044***	0.008***	-0.000	0.000
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_9	0.077***	0.004*	0.000	-0.002*	0.083***	0.005***	-0.000	-0.001
	(0.006)	(0.002)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
β_{10}	0.203***	0.017***	-0.002***	-0.002	0.204***	0.016***	-0.001	0.001
	(0.005)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	442,353	579,965	558,722	428,591	1,073,779	1,454,292	1,524,108	1,158,163
R^2	0.824	0.912	0.993	0.991	0.750	0.867	0.988	0.982
Within-R^2	0.039	0.010	0.000	0.001	0.039	0.012	0.000	0.001
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B21: Learning Parameters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.203***	0.190***	0.198***	0.246***	0.241***	0.251***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
λ_1	-2.225***	-2.223***	-2.157***	-1.800***	-1.929***	-1.854***
	(0.128)	(0.119)	(0.114)	(0.066)	(0.068)	(0.065)
λ_2	-2.362***	-2.313***	-2.265***	-1.881***	-2.012***	-1.951***
	(0.137)	(0.126)	(0.121)	(0.073)	(0.074)	(0.072)
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B22: Risk Aversion and Bias Including Zero Weights - Summary Statistics

	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.006	1.501
standard deviation	0.056	5.266
median	0.009	1.441
min	-0.431	-43.532
max	0.398	42.658
skewness	-1.111	0.954
kurtosis	13.375	13.727

Figure B12: Weighting Functions - Various Examples

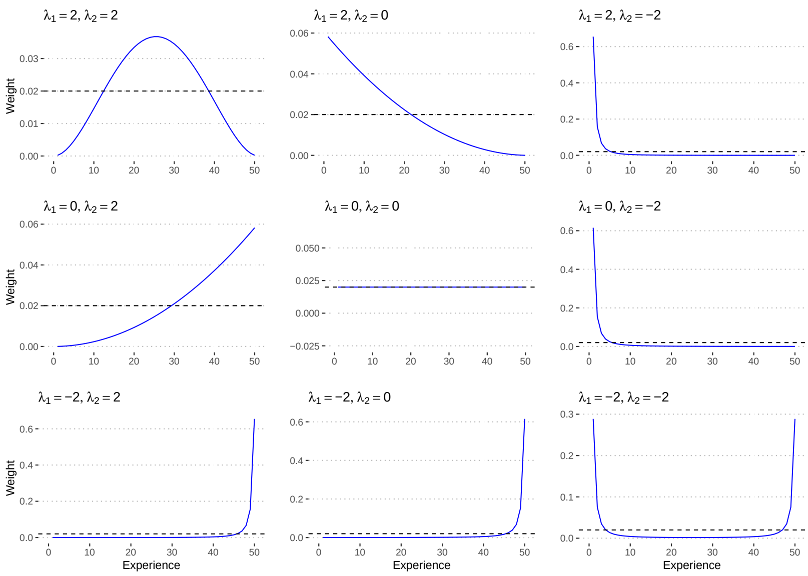
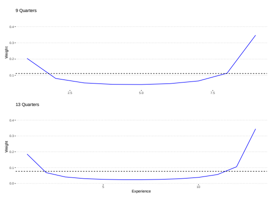
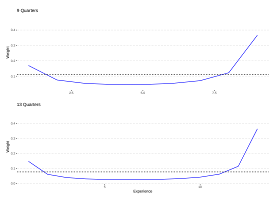


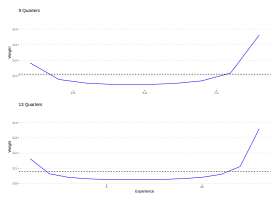
Figure B13: Estimated Weighting Functions - Manager-Time, Stock-Time Fixed Effects



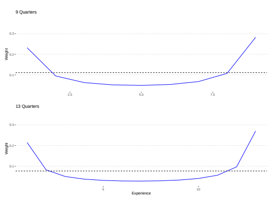
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^{*}, w_{i,j,t} > 0$



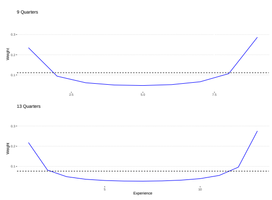
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^{*}, w_{i,j,t} > 0$



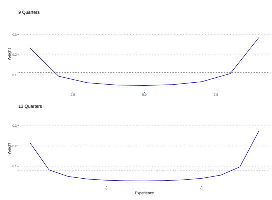
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^{*}, w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^{*},\ w_{i,j,t} \geq 0$

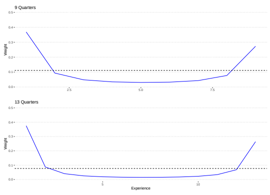


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^{*},\ w_{i,j,t} \geq 0$

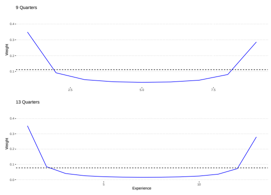


(f) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^{*},\ w_{i,j,t} \geq 0$

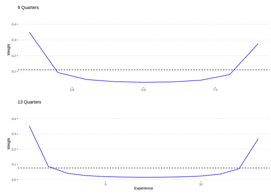
Figure B14: Estimated Weighting Functions - Manager-Time, Stock Fixed Effects



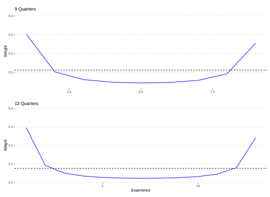
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



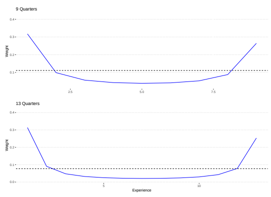
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



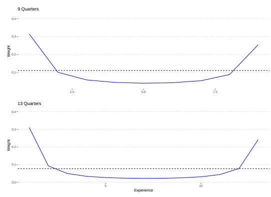
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

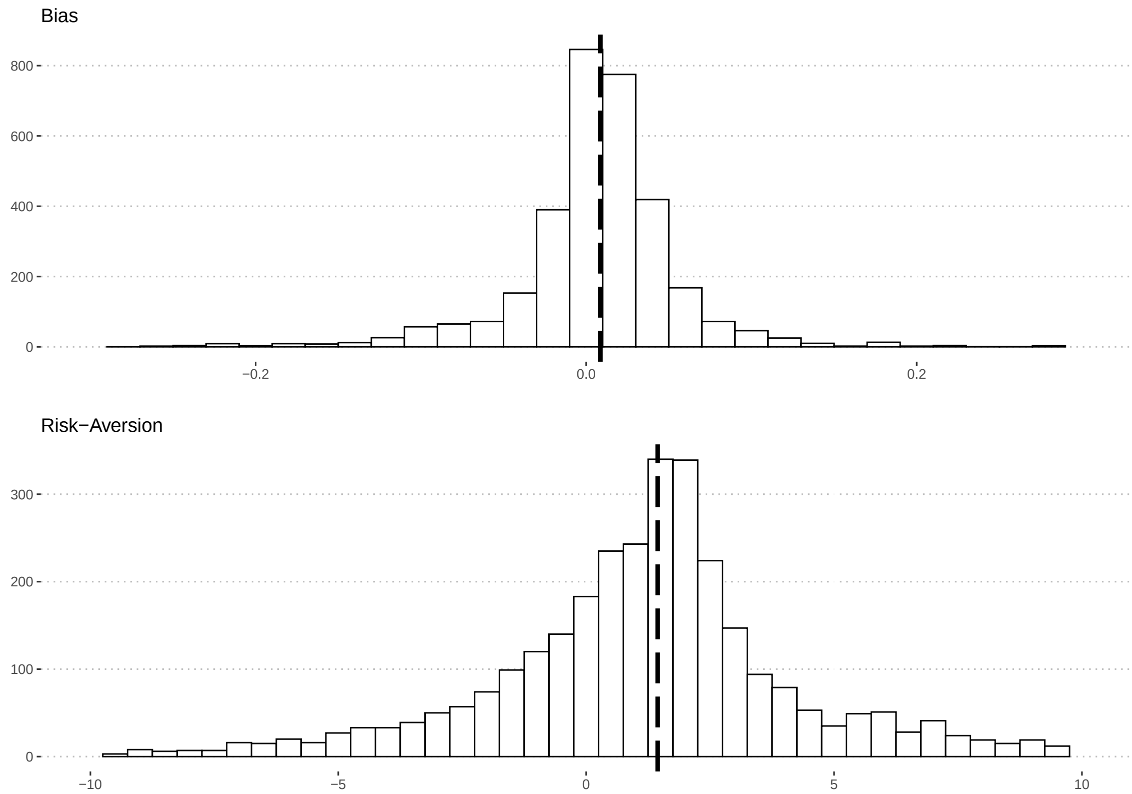


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$



(f) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

Figure B15: Bias and Risk Aversion Including Zero Weights



C. Appendix to Living on the Edge: the Salience of Property Taxes in the UK Housing Market

C.1 Variable Definition

Variable Name	Description
Price	Transaction price for the property as recorded by HM Land Registry
Council Tax	Amount of council tax payable per year
Band	Council tax band. One of: A, B, C, D, E, F, G, H
Year	Calendar year of the transaction
Month	Calendar month of the transaction
Size	Total floor area measured in squared meters
No. Rooms	Number of habitable rooms in the property as defined in the EPC
Property Type	One of: detached, semi-detached or terraced house and flat
Newly-built	Equals 1 if the property is newly-built
Leasehold	Equals 1 if the property is under a leasehold agreement
Energy Cost	Sum of the annual heating, hot water and lighting costs for the property
	One of very low, low, medium, high and very high expenditures
	Baseline = very low
CO_2 Emissions	CO_2 emissions in tonnes/year
	One of very low, low, medium, high and very high
	Baseline = very low
No. Lighting Outlets	Number of fixed lighting outlets in the property, standardised
Energy Rating	A-G energy rating fixed effects with A being the most efficient
Glazed Type	Indicates the type of glazing
	Various categories of single, double or triple glazing according to
	the British Fenestration Rating Council or manufacturer declaration
No. Storeys > 3	Equals 1 if the building has more than 3 storeys
Glazed Area	Estimate of total glazed area of the property

Continued on next page

Continued from previous page

	One of: Normal, Less than Normal, More than Normal
	Baseline = Normal
Fireplaces	Equals 1 if the property has open fireplaces
No. Extensions	Number of extensions added to the property
	One of: 0, 1, 2, 3, 4
Floor Height	Average storey height in metres
	One of: less than 2.3, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 or more
Built in	Age band when the building was constructed
	One of: before 1949, 1950-1982, 1983-2002, after 2003
Grid ID	An indicator for the grid square in which the property is located
Pair ID	An indicator for the pair of matched properties

C.2 Tables

Table C1: Evidence of Selection

The table shows the estimates of a simple regression of house prices on council tax amounts, namely: $p_{ibdt} = \beta\tau_{bdt} + \delta_{bt} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$ where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bt} are band-year fixed effects; and x_{ibdt} are controls. All columns include band-year and month fixed effects. All other variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-231.2***	-263.3***	-228.7***	-229.2***
	(71.8)	(86.4)	(78.0)	(78.3)
Size		2,233.7***	2,271.7***	2,270.8***
		(724.4)	(731.2)	(730.9)
Newly-built				14,054.3**
				(5,619.8)
Leasehold				-8,681.7
				(10,801.3)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Obs.	889,925	889,925	889,925	889,925
R^2	0.530	0.573	0.578	0.578
Within R^2	0.022	0.064	0.058	0.058

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C2: Evidence of Selection - Additional Controls

	(1)	(2)	(3)	(4)
Council Tax	-255.9***	-225.0**	-259.6***	-220.1**
	(85.6)	(80.0)	(88.2)	(78.2)
Size	2,747.2***	2,266.9***	2,534.4***	2,310.4***
	(911.3)	(734.5)	(780.3)	(784.4)
Energy Cost Low	-26,896.1*			-15,049.6**
	(13,515.0)			(7,107.5)
Energy Cost Medium	-47,312.9**			-24,385.3**
	(22,380.7)			(11,482.6)
Energy Cost High	-69,359.2**			-32,869.7**
	(30,818.3)			(15,291.8)
Energy Cost Very High	-94,269.8*			-39,075.4
	(45,563.6)			(22,987.4)
CO_2 Emisions Low	-17,677.3**			-14,199.3***
	(7,006.9)			(4,857.3)
CO_2 Emissions Medium	-26,558.8**			-23,257.6**
	(11,971.2)			(8,475.0)
CO_2 Emissions High	-36,052.5*			-31,559.2**
	(18,323.2)			(12,521.6)
CO_2 Emissions Very High	-32,523.2			-26,343.9
	(28,385.1)			(17,461.3)
No. Lighting Outlets	20,870.4***			19,659.8***
	(5,833.9)			(5,317.1)
No. Storeys > 3		-3,140.4		632.9
		(5,841.7)		(6,385.2)
Glazed Area Less than Normal		6,923.6		851.3
		(11,930.2)		(10,981.8)
Glazed Area More than Normal		16,669.1***		13,729.2***
		(3,337.1)		(3,490.2)
Fireplaces		42,454.0***		33,624.0***
		(9,985.6)		(9,114.6)
Newly-built			23,567.3***	29,368.6***
			(5,295.0)	(4,958.9)
Leasehold			24,601.4*	-13,104.3
			(13,060.7)	(11,681.6)

Continued on next page

Table C2 – *Continued from previous page*

	(1)	(2)	(3)	(4)
Built in 1950-1982			-43,868.7***	-29,435.6***
			(8,169.9)	(5,870.5)
Built in 1983-2002			-22,756.2**	-30,012.4***
			(9,533.2)	(8,919.4)
Built after 2003			-21,575.1	-31,925.1**
			(13,706.9)	(15,196.7)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Energy Rating	Yes	No	No	Yes
Glazed Type	Yes	No	No	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
No. Extensions	No	Yes	No	Yes
Floor Height	No	Yes	No	Yes
Obs.	889,925	889,925	889,925	889,925
R^2	0.566	0.580	0.564	0.583
Within R^2	0.095	0.059	0.092	0.063

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C3: Evidence of Selection - Median Price per Borough, Band, Year

The table shows the estimates of the following regression: $\varepsilon_{bdt}^{med} = \beta\tau_{bdt} + \delta_{bt} + \eta_{bdt}$, where ε_{bdt}^{med} is the median residual price of all houses in band b, Borough d at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bt} are band-year fixed effects. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-183.6***	-334.0***	-324.3***	-325.1***
	(56.4)	(84.7)	(83.5)	(83.1)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes
Obs.	5,014	5,014	5,014	5,014
R^2	0.804	0.501	0.503	0.500
Within R^2	0.055	0.122	0.117	0.118

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C4: Evidence of Selection - Average Price per Borough, Band, Year

The table shows the estimates of the following regression: $\bar{\varepsilon}_{bdt} = \beta\tau_{bdt} + \delta_{bt} + \eta_{bdt}$, where $\bar{\varepsilon}_{bdt}$ is the average residual price of all houses in band b, Borough d at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bt} are band-year fixed effects. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-195.6***	-368.4***	-358.6***	-358.9***
	(64.9)	(93.9)	(92.8)	(92.5)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes
Obs.	5,014	5,014	5,014	5,014
R^2	0.797	0.512	0.513	0.511
Within R^2	0.053	0.123	0.118	0.118

Two-way (Borough & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C5: Grid Regressions

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. All columns include band-grid ID-year and month fixed effects. The squares are constructed from a 50 $\times$ 50 grid of London. All other variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	50.3	12.6	13.4	14.3
	(50.9)	(48.0)	(45.3)	(44.7)
Size		4,626.9***	4,547.6***	4,537.0***
		(1,380.6)	(1,368.4)	(1,366.9)
Newly-built				33,398.5***
				(9,937.9)
Leasehold				-75,924.3**
				(27,874.0)
Fixed-effects				
Band $\times$ Grid ID $\times$ Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Obs.	71,734	71,734	71,734	71,734
R^2	0.696	0.771	0.773	0.773
Within R^2	0.000	0.103	0.010	0.101

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C6: Grid Regressions - Additional Controls

	(1)	(2)	(3)	(4)
Council Tax	7.98	15.0	9.19	17.5
	(42.3)	(45.4)	(40.1)	(43.4)
Size	5,855.9***	4,522.7***	5,318.7***	4,787.6***
	(1,585.1)	(1,366.3)	(1,353.6)	(1,548.8)
Energy Cost Low	-66,317.8**			-35,546.7**
	(23,418.5)			(14,809.1)
Energy Cost Medium	-108,665.0**			-57,508.8**
	(38,438.8)			(23,978.0)
Energy Cost High	-147,580.7**			-77,036.8**
	(52,657.2)			(33,410.5)
Energy Cost Very High	-195,178.8**			-106,670.2*
	(79,418.3)			(52,374.7)
CO_2 Emissions Low	-32,054.7***			-28,497.1***
	(10,727.2)			(9,893.5)
CO_2 Emissions Medium	-48,961.2**			-47,738.6***
	(17,139.3)			(16,123.8)
CO_2 Emissions High	-75,329.8**			-71,914.7***
	(27,138.1)			(24,295.9)
CO_2 Emissions Very High	-69,844.7			-66,123.0*
	(41,209.7)			(33,075.1)
No. Lighting Outlets	21,965.9**			19,370.8**
	(8,176.9)			(7,921.1)
No. Storeys > 3		-20,775.9***		-22,481.8***
		(6,704.8)		(7,545.2)
Glazed Area Less than Normal		-25,624.5		-18,393.5
		(18,311.0)		(17,084.0)
Glazed Area More than Normal		13,298.3		12,980.2
		(8,438.6)		(8,365.9)
Fireplaces		34,202.3***		32,533.3***
		(6,533.8)		(6,832.4)
Newly-built			23,232.4**	23,142.1*
			(9,272.9)	(11,929.6)
Leasehold			-64,925.4***	-81,662.8***
			(17,738.1)	(28,274.4)

Continued on next page

Table C6 – *Continued from previous page*

	(1)	(2)	(3)	(4)
Built in 1950-1982			-33,647.4***	-31,823.5***
			(8,513.9)	(9,833.6)
Built in 1983-2002			43,030.8**	5,862.8
			(19,392.1)	(10,375.6)
Built after 2003			37,169.2**	-1,989.2
			(16,898.5)	(15,825.0)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Energy Rating	Yes	No	No	Yes
Glazed Type	Yes	No	No	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
No. Extensions	No	Yes	No	Yes
Floor Height	No	Yes	No	Yes
Obs.	71,734	71,734	71,734	71,734
R^2	0.762	0.774	0.759	0.777
Within R^2	0.216	0.010	0.209	0.110

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C7: Grid Regressions for Different Grids

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. The grids divide London into 50 $\times$ 50, 100 $\times$ 100 and 150 $\times$ 150 squares in columns (1), (2) and (3), respectively. All columns include band-grid ID-year, month, number of rooms, property type, newly-built and leasehold fixed effects, as well as a control for the property size. All variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)
Council Tax	14.3	-16.2	28.2
	(44.7)	(58.4)	(32.4)
Size	4,537.0***	6,988.4***	7,737.4***
	(1,366.9)	(1,794.8)	(2,319.6)
Newly-built	33,398.5***	22,929.9	-28,536.5
	(9,937.9)	(20,993.7)	(24,742.0)
Leasehold	-75,924.3**	-82,738.9*	-151,551.3**
	(27,874.0)	(44,068.4)	(69,763.6)
Fixed-effects			
Band $\times$ Grid ID $\times$ Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
Obs.	71,734	21,446	6,954
R^2	0.773	0.792	0.827
Within R^2	0.101	0.139	0.154
Grid	50 $\times$ 50	100 $\times$ 100	150 $\times$ 150

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C8: Grid Regressions for Different Grids - Additional Controls

	(1)	(2)	(3)
Council Tax	17.5	-11.8	75.4**
	(43.4)	(62.0)	(33.1)
Size	4,787.6***	7,579.1***	7,516.7***
	(1,548.8)	(1,968.4)	(2,014.1)
Newly-built	23,142.1*	23,012.5	-17,873.0
	(11,929.6)	(23,337.2)	(11,886.9)
Leasehold	-81,662.8***	-100,958.8*	-180,856.2*
	(28,274.4)	(48,348.3)	(92,255.8)
Built in 1950-1982	-31,823.5***	-36,770.1**	-46,677.3*
	(9,833.6)	(13,202.8)	(22,562.5)
Built in 1983-2002	5,862.8	26,842.2	-20,068.4
	(10,375.6)	(22,698.3)	(26,632.8)
Built after 2003	-1,989.2	-30,612.1	-68,073.3
	(15,825.0)	(29,315.1)	(66,304.7)
No. Storeys > 3	-22,481.8***	-19,920.1**	-10,496.4
	(7,545.2)	(9,293.2)	(10,564.8)
Glazed Area Less than Normal	-18,393.5	-37,901.5	41,021.3
	(17,084.0)	(25,339.4)	(69,940.7)
Glazed Area More than Normal	12,980.2	4,587.5	-102,420.2
	(8,365.9)	(19,443.3)	(63,902.4)
Fireplaces	32,533.3***	41,107.7***	49,004.1***
	(6,832.4)	(12,251.9)	(16,591.1)
Energy Cost Low	-35,546.7**	-55,601.8**	-62,161.9***
	(14,809.1)	(20,340.8)	(16,726.8)
Energy Cost Medium	-57,508.8**	-93,685.8***	-85,741.3***
	(23,978.0)	(31,367.3)	(23,759.2)
Energy Cost High	-77,036.8**	-141,100.4***	-161,362.9***
	(33,410.5)	(46,540.7)	(41,048.0)
Energy Cost Very High	-106,670.2*	-170,909.0**	-189,343.0**
	(52,374.7)	(66,032.5)	(72,613.7)
CO_2 Emissions Low	-28,497.1***	-46,607.7***	-43,592.6***
	(9,893.5)	(14,538.3)	(13,312.9)
CO_2 Emissions Medium	-47,738.6***	-73,467.2***	-96,311.1**
	(16,123.8)	(23,582.7)	(43,159.3)

Continued on next page

Table C8 – *Continued from previous page*

	(1)	(2)	(3)
CO_2 Emissions High	-71,914.7***	-104,329.4***	-141,727.0**
	(24,295.9)	(32,544.5)	(58,712.5)
CO_2 Emissions Very High	-66,123.0*	-133,060.2***	-150,608.5*
	(33,075.1)	(44,218.2)	(77,511.2)
No. Lighting Outlets	19,370.8**	22,306.9	53,060.6*
	(7,921.1)	(16,072.2)	(30,324.9)
Fixed-effects			
Band × Grid ID × Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
No. Extensions	Yes	Yes	Yes
Floor Height	Yes	Yes	Yes
Energy Rating	Yes	Yes	Yes
Glazed Type	Yes	Yes	Yes
Obs.	71,734	21,446	6,954
R^2	0.777	0.798	0.846
Within R^2	0.110	0.150	0.165
Grid	50 × 50	100 × 100	150 × 150

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C9: Grid Regressions - Without Stamp Duty Notches

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. All columns include band-grid ID-year, month, number of rooms, property type, newly-built and leasehold fixed effects, as well as a control for property size. The squares are constructed from a 50×50 grid of London. Column (1) excludes properties sold at a price between £240, 000 and £270, 000; column (2) properties sold for between £490, 000 and £520, 000; and column (3) excludes both properties sold in the £240, 000 - £270, 000 and £490, 000 - £520, 000 price range. All variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)
Council Tax	16.1	16.7	18.8
	(46.9)	(45.3)	(47.7)
Size	4,715.8***	4,586.7***	4,765.9***
	(1,446.6)	(1,403.6)	(1,487.4)
Newly-built	37,062.1***	30,619.3***	33,964.1***
	(9,998.4)	(9,694.0)	(9,549.4)
Leasehold	-80,083.0**	-75,897.2**	-80,141.1**
	(30,142.0)	(28,682.7)	(31,002.5)
Fixed-effects			
Band × Grid ID × Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
Obs.	65,328	70,012	63,606
R^2	0.775	0.776	0.779
Within R^2	0.105	0.102	0.106
$p \notin$	[240k-270k]	[490k-520k]	[240k-270k] & [490k-520k]

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C10: Grid Regressions - Median Price per Borough, Band, Grid, Year

The table shows the estimates of the following regression: $\varepsilon^{med}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt}$, where ε^{med}_{bdgt} is the median residual price of all houses in band b, Borough d, grid square g at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bgt} are band-grid ID-year fixed effects. The squares are constructed from a 50×50 grid of London. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the grid ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	92.1*	15.4	19.9	19.1
	(50.8)	(35.4)	(36.3)	(36.6)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Fit statistics				
Obs.	19,377	19,377	19,377	19,377
R^2	0.866	0.833	0.825	0.823
Within R^2	0.006	0.000	0.000	0.000
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C11: Grid Regressions - Average Price per Borough, Band, Grid, Year

The table shows the estimates of the following regression: $\bar{\varepsilon}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt}$, where $\bar{\varepsilon}_{bdgt}$ is the average residual price of all houses in band b, Borough d, grid square g at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bgt} are band-grid ID-year fixed effects. The squares are constructed from a 50 × 50 grid of London. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the grid ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	104.5**	23.6	28.2	26.6
	(48.0)	(32.7)	(33.5)	(33.9)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Obs.	19,377	19,377	19,377	19,377
R^2	0.875	0.835	0.827	0.825
Within R^2	0.007	0.001	0.001	0.001
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C12: Matching Regressions - Euclidean Distance

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. The closest match for each property is chosen as the one minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$. The vectors x_i and x_j in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	53.8**	12.9	50.7**	9.00
	(23.4)	(18.3)	(23.8)	(18.8)
Size		3,770.6***		3,750.2***
		(763.8)		(734.2)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	115,224	115,224	114,646	114,646
Unique Transaction IDs	71,578	71,578	71,656	71,656
R^2	0.799	0.836	0.796	0.834
Within R^2	0.001	0.042	0.001	0.042
Distance	Euclidean 1	Euclidean 1	Euclidean 2	Euclidean 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C13: Matching Regressions - Linear Distance

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. The closest match for each property is chosen as the one minimising the following distance: $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}|$, where $\hat{p}_{it}$ and $\hat{p}_{jt}$ are model-predicted prices for two matched property transactions i and j based on a linear model: $p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$. The vectors x_{it} and x_{jt} in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	56.8**	15.3	55.7**	14.6
	(23.4)	(18.1)	(23.7)	(18.7)
Size		3,879.2***		3,809.8***
		(778.8)		(762.1)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	114,904	114,904	113,854	113,854
Unique Transaction IDs	71,588	71,588	71,649	71,649
R^2	0.799	0.837	0.798	0.835
Within R^2	0.001	0.045	0.001	0.043
Distance	Linear 1	Linear 1	Linear 2	Linear 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C14: Matching Regressions - Linear Distance Less than 30% of Predicted Prices

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. Each house i is matched to all possible candidates j that satisfy the following constraint: $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}| < 0.3 \times \max\{\hat{p}_{it}, \hat{p}_{jt}\}$, where $\hat{p}_{it}$ and $\hat{p}_{jt}$ are model-predicted prices for two matched property transactions i and j based on a linear model: $p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$. The vectors x_{it} and x_{jt} in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-8.19	-5.24	-7.65	-8.14
	(10.1)	(9.68)	(11.0)	(10.3)
Size		3,980.1***		3,982.4***
		(295.8)		(349.3)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	175,639	175,639	167,704	167,704
Unique Transaction IDs	59,722	59,722	58,917	58,917
R^2	0.871	0.875	0.855	0.859
Within R^2	0.000	0.017	0.000	0.018
Distance	Linear 1	Linear 1	Linear 2	Linear 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C15: Model-averaged Posterior Distributions for the Council Tax Incidence

The table displays 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 99% quantiles, the modal and mean values of the average posterior distribution for the council tax incidence obtained by using the estimates from Tables C5-C9 and C12-C14. The last column reports the attenuation factor γ computed as the ratio of the posterior and prior median. Each row refers to a different choice of prior.

Prior	1%	5%	10%	25%	50%	75%	90%	95%	99%	mode	mean	γ
$\mathcal{N}(-150, 50^2)$	-143.50	-110.66	-93.21	-61.85	-22.87	-1.98	18.67	31.04	51.90	-12.08	-31.85	0.15
$\mathcal{N}(-100, 50^2)$	-116.75	-85.88	-69.23	-39.43	-12.79	7.60	29.51	41.86	62.81	-9.86	-16.81	0.13
$\mathcal{N}(-50, 50^2)$	-90.71	-61.45	-45.54	-20.99	-2.17	20.33	42.09	54.20	75.25	-6.78	-1.76	0.04
$\mathcal{N}(-150, 75^2)$	-126.67	-87.78	-67.60	-32.92	-7.49	16.49	41.31	54.86	78.03	-8.24	-10.46	0.05
$\mathcal{N}(-50, 25^2)$	-82.09	-64.43	-54.54	-36.79	-18.64	-4.40	9.15	17.64	32.93	-13.86	-20.87	0.37

C.3 Figures

Figure C1: A Typical Border

The figure shows an example of a border between two Boroughs in London. Houses on the left side of the West Eaton Place road belong to the Borough of Kensington and Chelsea and have an annual council tax bill of £2, 279, while houses on the right side belong to the Borough of Westminster and have an annual council tax bill of only £1, 421.

Figure C2: Time Series of Council Taxes

The figure reports the time series of council tax amounts payable across Boroughs. Each panel refers to a different band, while the lines in each panel represent different Boroughs.

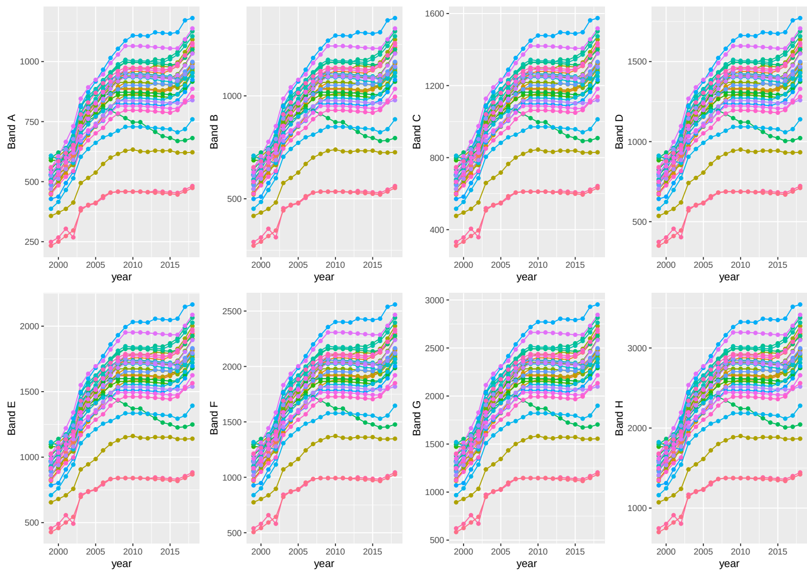


Figure C3: Histogram of Property Prices in London

The figure presents a histogram of the distribution of house transaction prices in London. The distribution is truncated at £1, 500, 000.

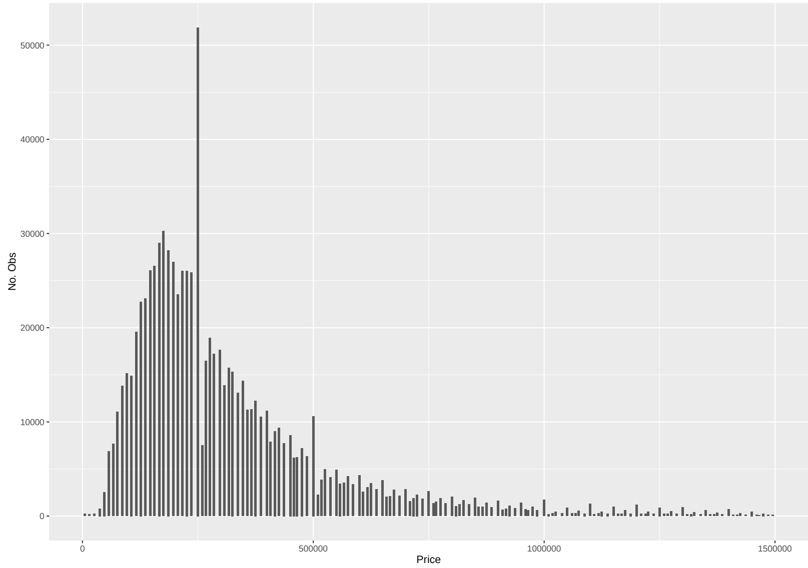


Figure C4: Bunching at Stamp Duty notches

The figure presents a histogram of the distribution of house transaction prices in London around stamp duty notches. Panel (A) refers to the notch at £250, 000, while panel (B) at £500, 000.

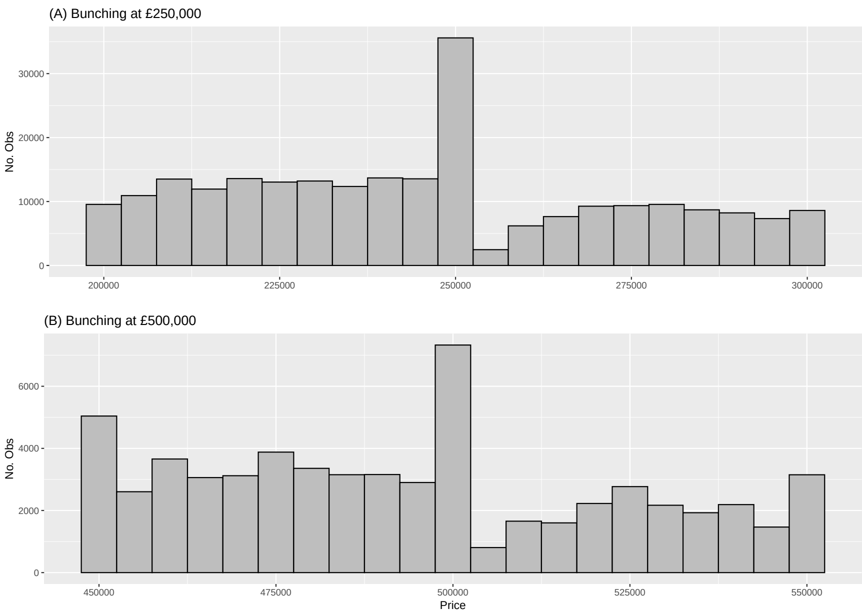


Figure C5: Histogram of Prices by Band

The figure presents a histogram of the distribution of house transaction prices in London per band. Each panel refers to properties belonging to different bands. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

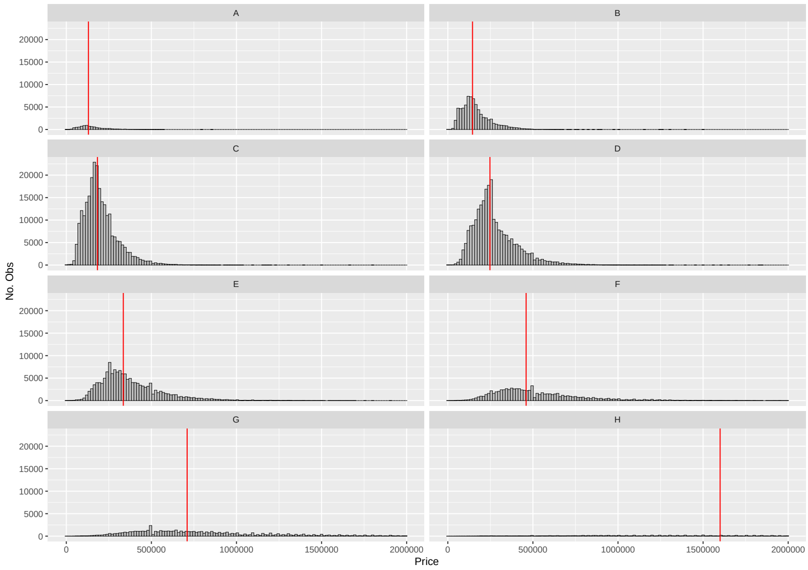


Figure C6: Histogram of Prices by Property Type

The figure presents a histogram of the distribution of house transaction prices in London by property type. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

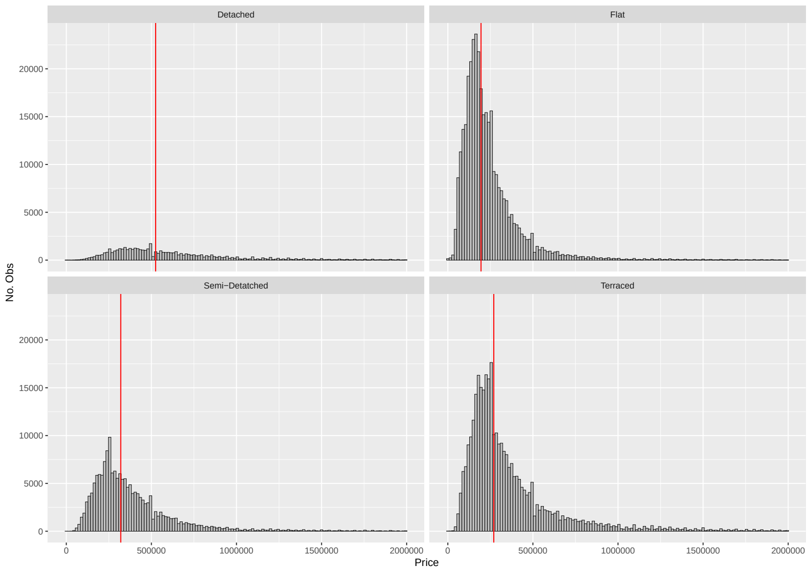


Figure C7: Histogram of Prices by Number of Rooms

The figure presents a histogram of the distribution of house transaction prices in London by number of rooms. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

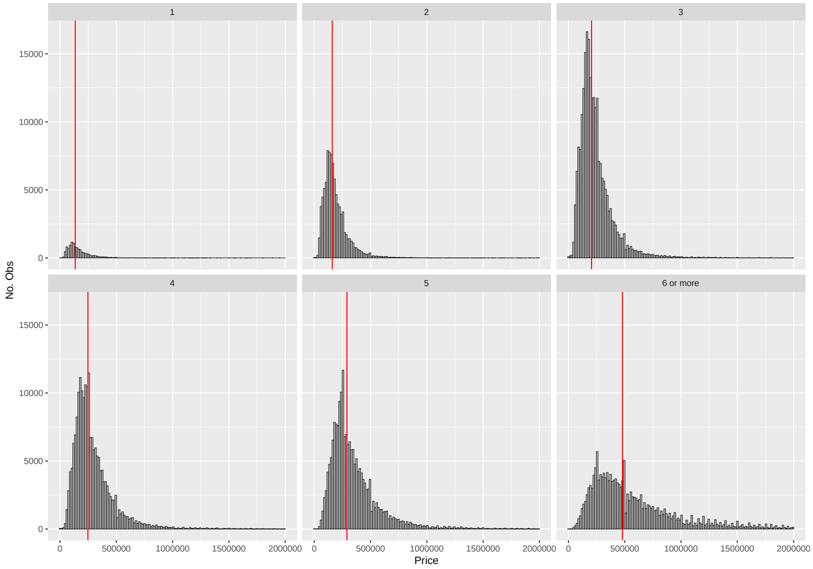


Figure C8: Histogram of Prices by Age

The figure presents a histogram of the distribution of house transaction prices in London by age. The top panel reports the histogram of prices for newly-built properties, while the bottom for established residential buildings. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

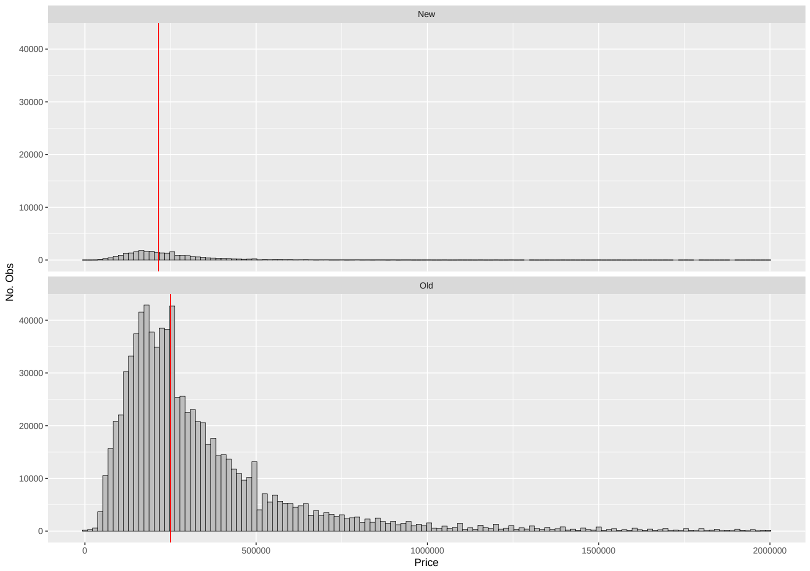


Figure C9: Histogram of Prices by Year of Construction

The figure presents a histogram of the distribution of house transaction prices in London by year of construction. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

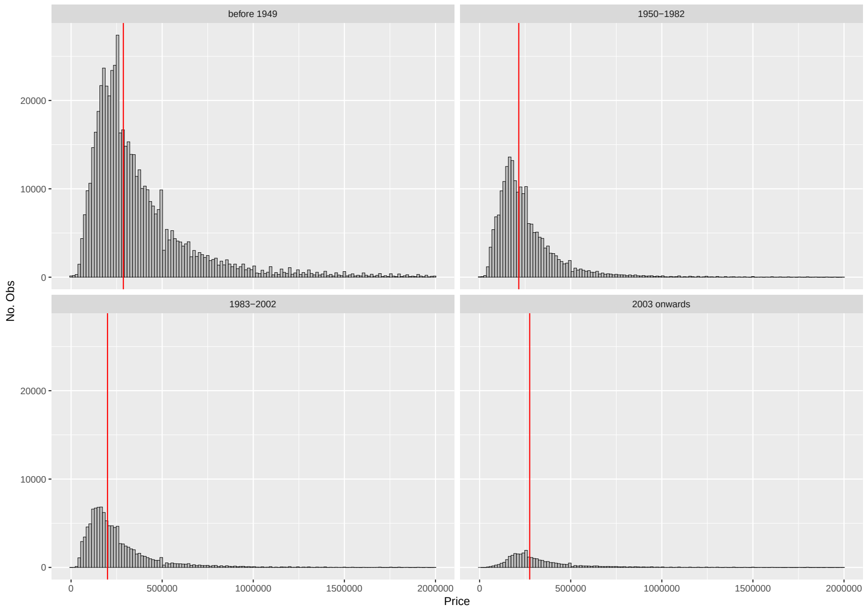


Figure C10: Histogram of Prices by Duration

The figure presents a histogram of the distribution of house transaction prices in London by tenure duration. The top panel reports the histogram of prices for freehold properties, while the bottom for leasehold properties. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

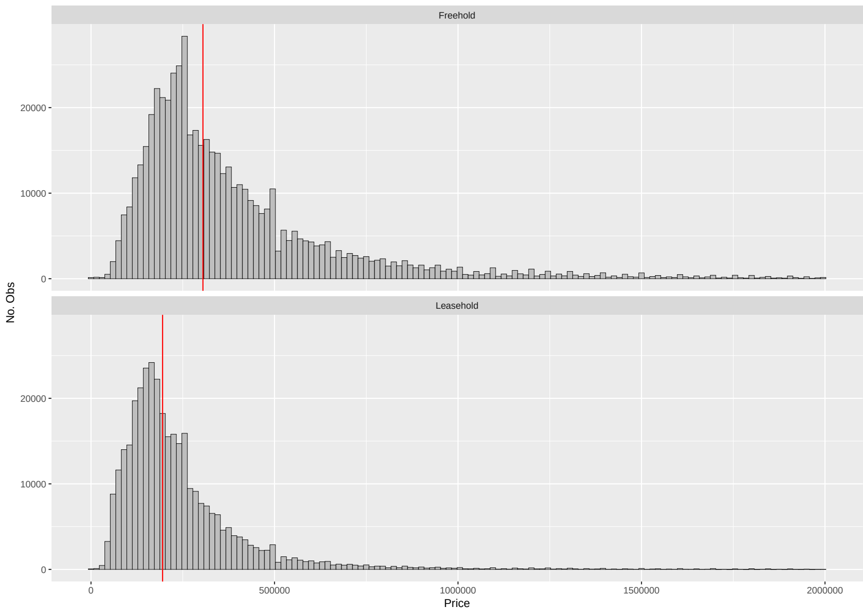
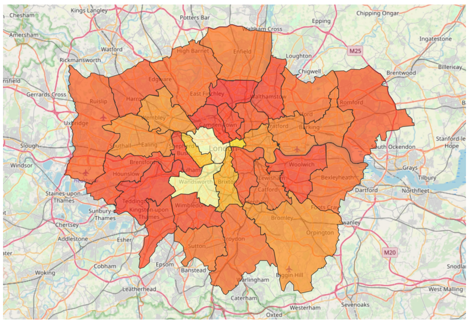
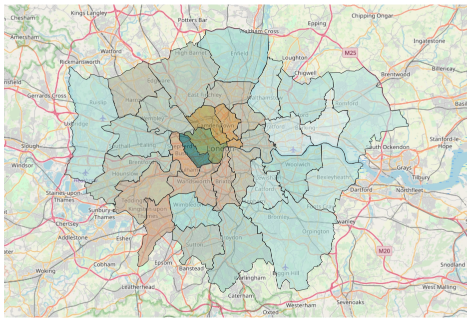


Figure C11: Council Taxes and House Prices

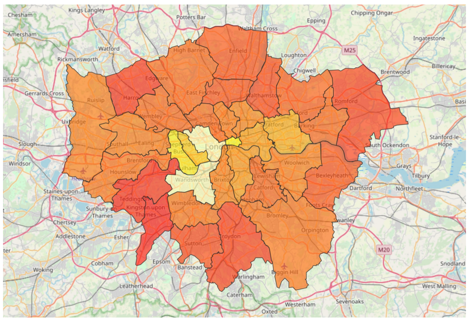
The maps show the distribution of council tax payable for properties in band D for each London Borough, along with the respective distribution of house prices in 2000 and 2018.



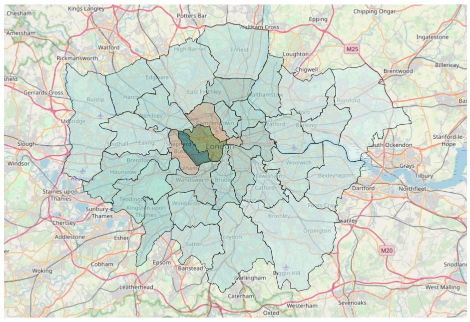
(a) Council Taxes in 2000



(b) House Prices in 2000



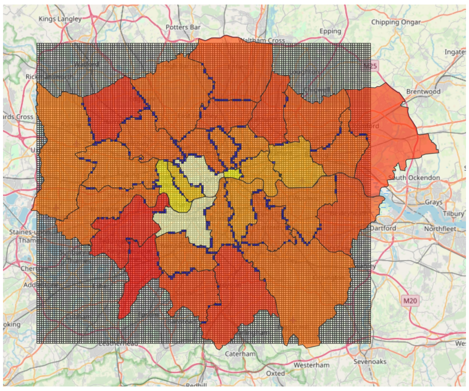
(c) Council Taxes in 2018



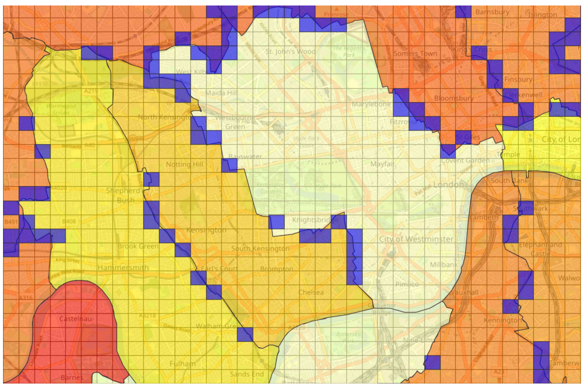
(d) House Prices in 2018

Figure C12: Grids

The maps depict our first identification strategy of dividing London in a grid of equally sized squares. Panel C12a shows a grid of 150 × 150 squares superposed on the map of the city; Panel C12b shows an enlargement of the central Boroughs. The blue squares denote areas which contain at least two similar properties located on opposite sides of a border.



(a) Grid



(b) Enlargement of the Centre

Figure C13: Distribution of Distances for the Grid Regressions

The figure depicts histograms for the distribution of distances between houses on opposite sides of a border that are used in our grid regressions. We report the distributions for three different grids, namely grids where we have divided London in 50 × 50 squares, 100 × 100 and, finally, 150 × 150. For each histogram we report the approximate size of the square sides in meters.

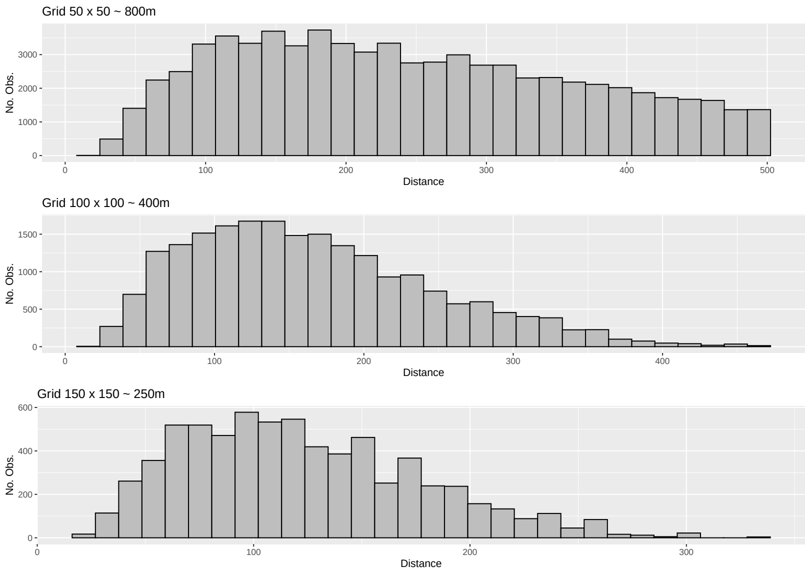


Figure C14: Model-implied Incidence

The figure plots the relationship between tax incidence on house prices and discount rates, where the discount rate is defined as $r + k$ as in Section 3.4. The upper panel shows the incidence of the stamp duty, while the bottom panel the incidence of the council tax.

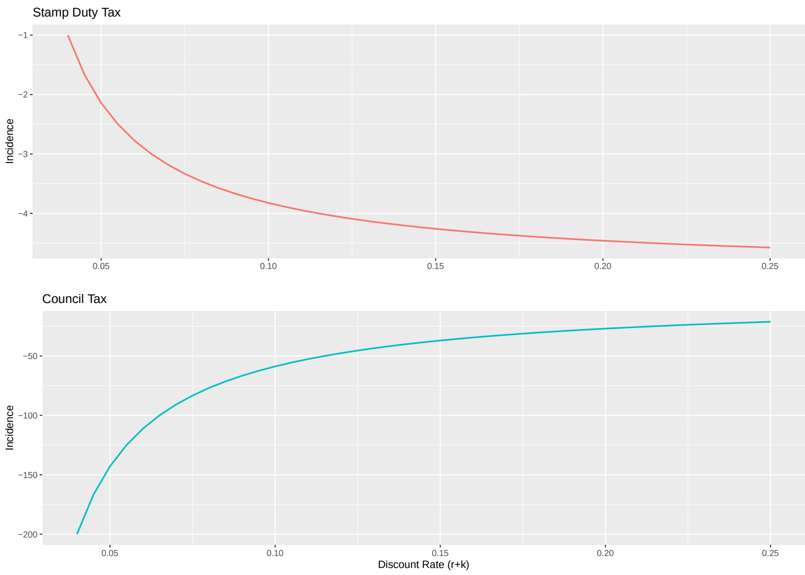
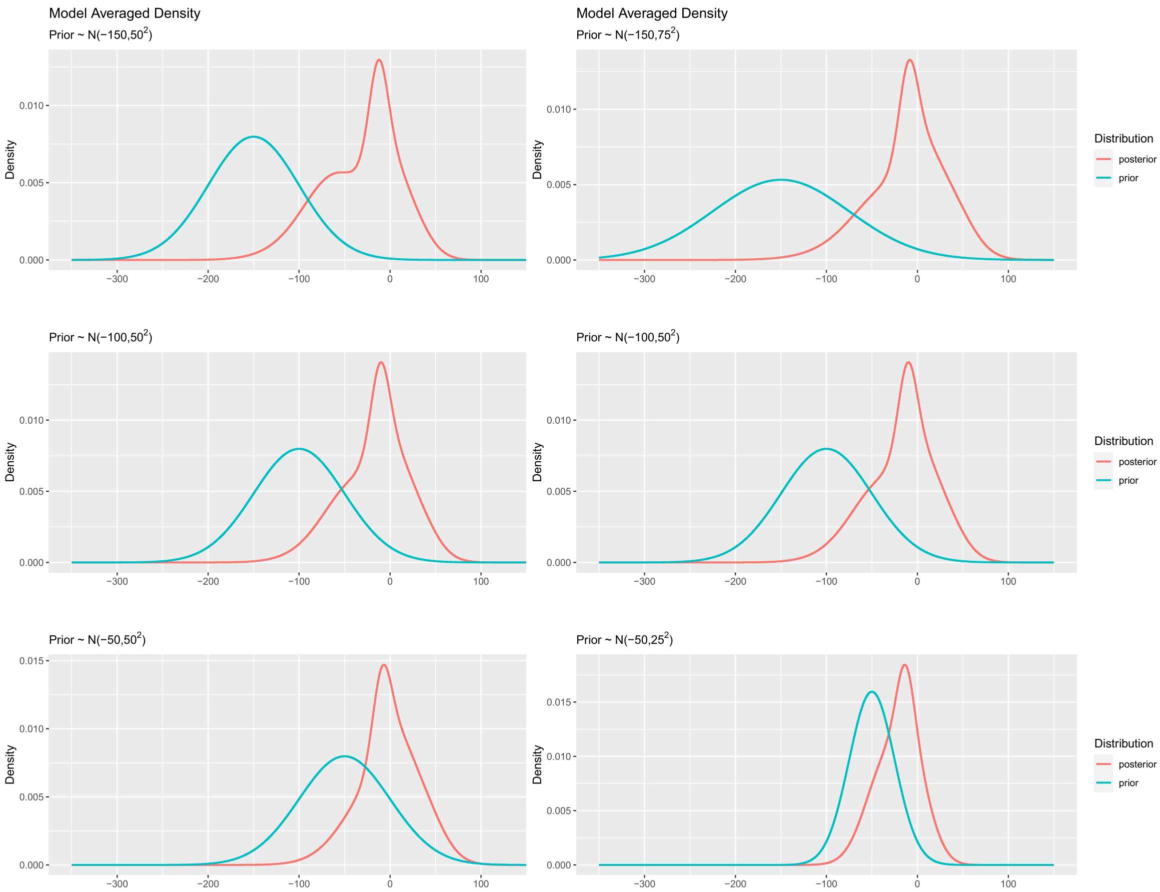


Figure C15: Model-averaged Estimate of the Posterior Council Tax Incidence

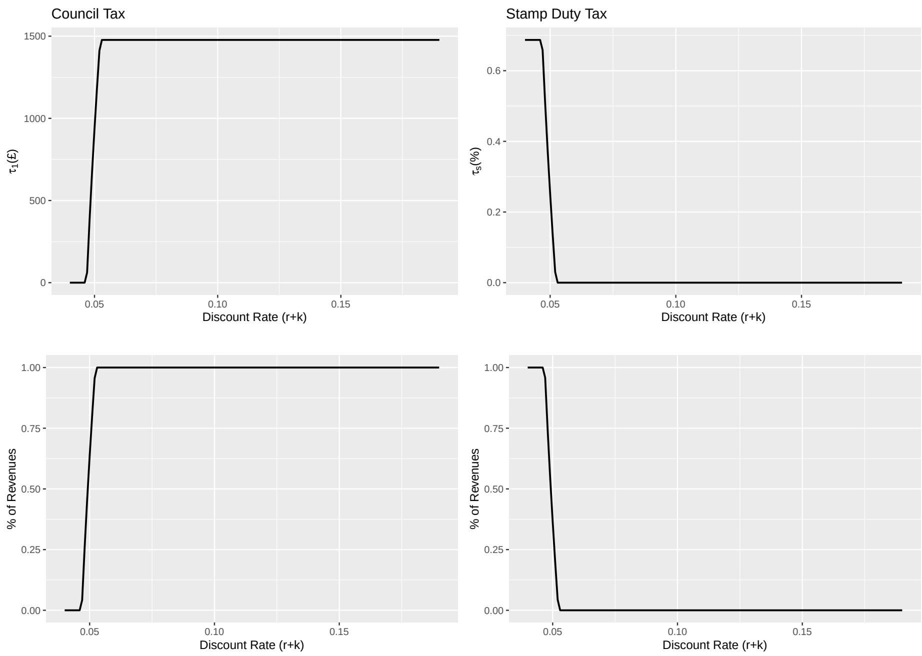
The figure plots the density of the council tax incidence obtained by taking the model-average of the posteriors as described in Sections 3.4.1 and C.4. The priors are normally distributed $\mathcal{N}(b_0, \sigma_0^2)$ in all figures. In panel (a) the priors have constant standard deviation $\sigma_0 = 50$ and varying means of $b_0 = -150, -100, -50$, respectively. In panel (b) the standard deviation of the priors is proportional to the mean, i.e., $\sigma_0 = |b_0|/2$.



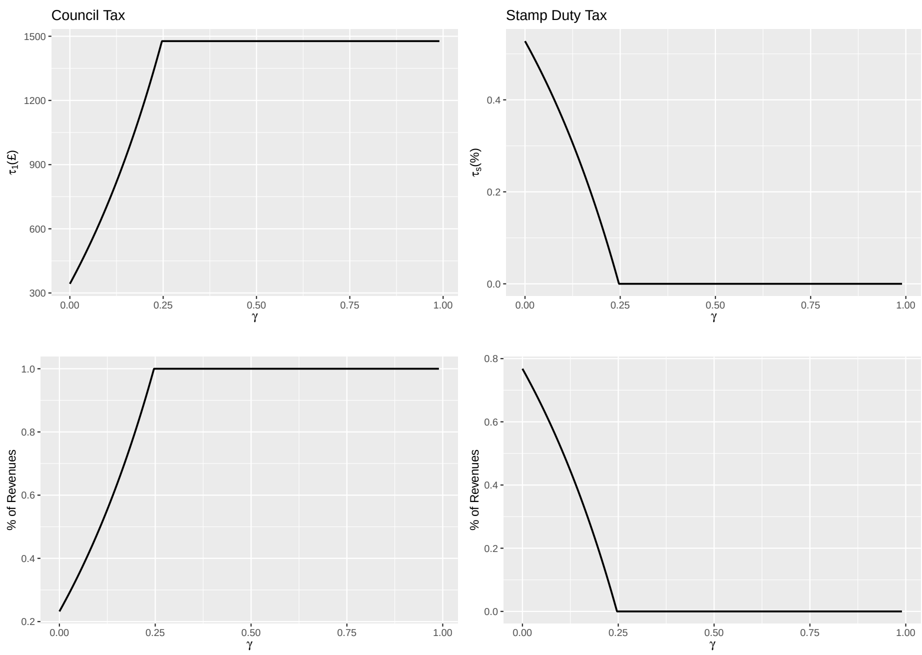
(a) Constant Variance (b) Proportional Variance

Figure C16: Optimal Tax Policy

The figure plots the optimal mix of stamp duty and council tax the Government should choose to maximise the utility of buyers and maintain revenue-neutrality. Panel (a) displays the variables as a function of the discount rate $r + k$, while panel (b) as a function of the attenuation parameter γ. The top plots of each panel show the optimal amount of council tax in £ and stamp duty tax as percentage of house price, respectively. The bottom plots provide the relative percentages of revenue raised through council and stamp duty tax, respectively. In the upper panel we calibrate the parameters as follows: $\alpha = 0.8$, $g = \tilde{g} = 3.5\%$, $\eta_S = 0.5$, $\beta = 0.99$, $\gamma = 0.15$; in the bottom panel: $\alpha = 0.8$, $g = \tilde{g} = 3.5\%$, $\eta_S = 0.5$, $\beta = 0.99$, $r + k = 5\%$.



(a) Optimal Taxes as a Function of $r + k$



(b) Optimal Taxes as a Function of γ

C.4 Computation of the Model-averaged Posterior Incidence of Council Tax

In Section 3.3 we estimate models of the type:

$$y = X^m \beta^m + \varepsilon^m \tag{C.1}$$

where $\varepsilon^m|m \sim \mathcal{N}(0, \Omega^m)$, with Ω^m being the population covariance matrix of the errors under model m. We partition the parameters as $\beta^m = (\beta_0, \beta^m_{-0})$, where $\beta^m_{-0} = (\beta^m_1, \beta^m_2, ...)$ and β_0 is the parameter of interest. We then make the (strong) simplifying assumption that Ω^m is known and assume that the prior distribution of the parameters is: $\beta^m|m \sim \mathcal{N}(b^m, \Sigma^m)$. We also assume that the marginal prior distribution of the parameter of interest is common across models, i.e., $p(\beta_0|m) = p(\beta_0) = \mathcal{N}(b_0, \sigma_0^2)$. It follows that the posterior is: $\beta^m|y, m \sim \mathcal{N}(((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1}X^m)^{-1}(X^{m\prime}(\Omega^m)^{-1}y + (\Sigma^m)^{-1}b^m), ((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1}X^m)^{-1})$. We then proceed by making the following approximations:

$$((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1}X^m)^{-1}_{[1,1]} \approx (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1} \tag{C.2}$$

$$\begin{aligned}(((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1}X^m)^{-1}(X^{m\prime}(\Omega^m)^{-1}y + (\Sigma^m)^{-1}b^m))_{[1]} \approx \\ (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}(\widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]}\hat{\beta}^m_0 + \sigma_0^{-2}b_0)\end{aligned} \tag{C.3}$$

where $A_{[i,j]}$ and $a_{[i]}$ indicate the ij-th element of matrix A and the i-th element of vector a, respectively. This leads, therefore, to the following approximate posterior distribution for the parameter of interest:

$$\begin{aligned}&p(\beta_0|y, m) = \\ &\mathcal{N}\left((\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}(\widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]}\hat{\beta}^m_0 + \sigma_0^{-2}b_0), (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}\right)\end{aligned} \tag{C.4}$$

After having obtained the posterior distribution for β_0 for each model we average using a flat prior across models to obtain the final density $p(\beta_0|y) = \frac{1}{M}\sum_{m=1}^{M} p(\beta_0|y, m)$.

Returning to the choice of prior distribution for the parameter of interest, we are guided by the model-implied incidence from Section 3.4. We calibrate the following parameters: $g = 0.035$, $\tilde{g} = 0.035$, $r = 0.04$ and $\alpha = 0.8$[1]. Given these values we pick three different means for the prior distribution to match the range of incidence of the stamp duty tax obtained in Best and Kleven (2018), namely, $b_0 = -150, -100, -50$, which roughly correspond to stamp duty incidences of: $\frac{dp}{d\tau_S} = -2, -3, -4$. We choose the standard deviations of the prior to be equal to $\sigma_0 = 50$ or $\sigma_0 = \frac{|b_0|}{2}$ to obtain five prior distributions.

[1]The parameters r and $\tilde{g}$ are consistent with the in-sample average mortgage rate and growth rate of council taxes in the UK, respectively; α is consistent with a downpayment of 20% which is common in the UK. We use a conservative expected growth rate of house prices of 3.5% compared to the in-sample average of 7.3%.