The London School of Economics and Political Science

Essays in Macro-Finance

Akash Raja

A thesis submitted to the Department of Economics
for the degree of Doctor of Philosophy

August 2, 2023

Acknowledgements

Doing a PhD can be both exciting and terrifying at the same time, and there are so many people to thank for supporting me along the way. First, to my incredible advisors - Ethan Ilzetzki, Ben Moll, Cameron Peng, and Ricardo Reis. I will always be grateful for the time you have spent guiding me through the research process, especially during the multiple times when I struggled. You have always encouraged me to think deeply about questions and pushed me to reach my potential. I thank all the faculty I have interacted with during my time at the LSE, in particular Francesco Caselli, Jonathon Hazell, Wouter den Haan, Dmitry Mukhin, and Maarten de Ridder. The supportive and constructive environment at the LSE has been key to making me the researcher I am today. I thank Dirk Jenter and Mark Schankerman for all their support as placement directors during my job market, as well as all the support staff at the LSE.

During my PhD, I was fortunate to be able to undertake internships at the Bank of England and Norges Bank, and learned so much from these experiences and my wonderful coauthors, namely David Aikman, Pietro Biroli, Teodora Boneva, Jonathan Bridges, Sigurd Mølster Galaasen, Sinem Hacioğlu Hoke, Cian O'Neill, and Christopher Rauh. Big thanks also to the University of Chicago, in particular Fernando Alvarez, Lars Peter Hansen, and Esteban Rossi-Hansberg, for inviting me to visit and meet so many amazing students and faculty.

To all my PhD colleagues, I will cherish the memories we have had going through the process together. Thank you for taking the time to listen to my practice presentations, for helping me whenever I was stuck, and most importantly for joining me for beers at the George IV pub. Special thanks go to Rasif Alakbarov, Daniel Albuquerque, Philipp Barteska, Fabio Bertolotti, Alix Bonargent, Thomas Brzustowski, Gaia Dossi, Shadi Farahzadi, Xitong Hui, Tomer Ifergane, Edoardo Leonardi, Junyi Liao, Will Matcham, Virginia Minni, Adi Soenarjo, Bilal Tabti, and Heidi Thysen.

To my friends from Cambridge - Aaran, Amar, Chengran, Francesca, James, Jamie, and

Muqeet. You always cheer me up when I am stressed and never fail to make me laugh. We have such great memories together from our undergraduate studies, and I'm so grateful that our friendship is so strong even today.

To my school friends - Gandhi, Jaymin, Parth, Pattni, Raees, Sheel, and Suraj. We have grown up together, and are truly a family of brothers. Thank you for teaching me to celebrate the wins, and for cheering me on every step of the way.

And last, but by no means least - to my parents, Asha and Kishor, and my sister, Sonam. You have always supported and believed in me, even when I didn't believe in myself. No words can express how grateful I am for everything you have given me, and I know that this would never have been possible without you. We did it.

Declaration

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 authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 35,218 words.

Statement of conjoint work

I confirm that Chapter 1 was jointly co-authored with Sigurd Mølster Galaasen, and I contributed 50% of this work. Chapter 4 was jointly co-authored with David Aikman, Jonathan Bridges, Sinem Hacioğlu Hoke, and Cian O'Neill, and I contributed 20% of this work.

Statement of inclusion of previous work

I confirm that Chapter 2 was the result of previous study for the MRes in Economics award I undertook at the London School of Economics and Political Science.

Abstract

This thesis consists of four chapters. The first chapter (co-authored with Sigurd Galaasen) studies the stock market entry and exit decisions of retail investors. Using Norwegian administrative data, we show that transitory spells in the stock market are common. A workhorse portfolio choice model requires sizable per-period participation costs to generate these patterns. We propose a theory of experience effects, whereby agents form beliefs over future stock returns based on their own realised returns. This model can explain the short-term dynamics in participation without requiring high costs.

The second chapter analyses how banks respond to capital regulation using data on bank-specific requirements in the UK. I find that actual capital ratios adjust following changes in requirements, though the pass-through is incomplete. The adjustment occurs primarily through capital accumulation and the risk composition of the asset portfolio. I find that the reaction of banks depends on the sign of the regulatory change and has changed since the financial crisis.

The third chapter exploits the implementation of the Basel I capital regulations in the US to study how bank capital affects lending. Using a difference-in-differences strategy, I show that undercapitalised banks reduced the size of their balance sheet to adjust to the new capital requirements. This decline in total assets is concentrated in loans, particularly commercial & industrial and non-residential real estate loans, with residential mortgage lending remaining unaffected.

The final chapter (co-authored with David Aikman, Jonathan Bridges, Sinem Hacioğlu Hoke, and Cian O'Neill) explores the relationship between financial vulnerabilities and downside risks to economic growth. Using quantile regressions applied to cross-country panel data, we show that credit and property price booms and wide current account deficits increase downside risks to GDP growth in the medium term. Such risks can be partially mitigated by increasing banking sector capitalisation.

Contents

List of Figures

List of Tables

Chapter 1

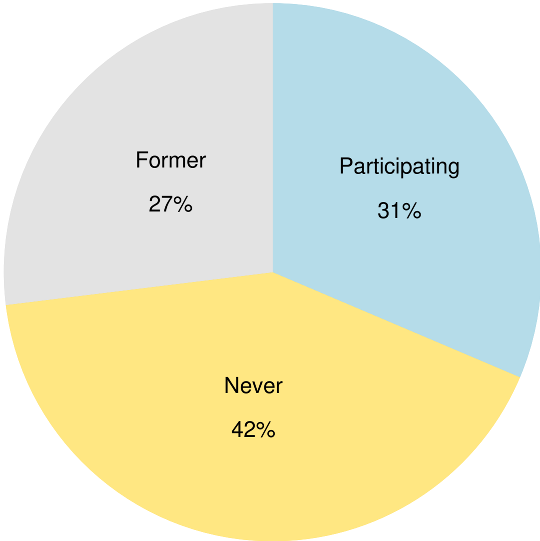
The Dynamics of Stock Market Participation

1 Introduction

Despite the large average return on equities relative to bonds, many households choose not to invest in the stock market (Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995; Campbell, 2006). While the literature has devoted significant attention to explaining why the aggregate participation rate lies below 100%, much less is known about the movements in and out of the stock market by individual investors. The conventional view is that people either never or always invest. However, the data indicate that a sizable proportion of nonparticipants have invested in stocks at some point in the past (Figure 1.1). Exploring the decision to enter into or exit from the stock market is of first-order importance because portfolio choices matter for wealth accumulation (Benhabib et al., 2011; Gabaix et al., 2016; Xavier, 2021). Furthermore, analyzing these transitions can help distinguish between the wide range of existing theories of participation, given that different models have opposing predictions for such movements. In this paper, we shed light on the dynamics of stock market participation by uncovering novel facts pertaining to exit and reentry at the individual level using detailed Norwegian administrative data, and assess the implications of our findings for theories of participation.

Panel data on wealth holdings over a long time dimension are essential to investigate changes in participation status. We exploit Norwegian administrative tax records to overcome this challenge. As Norway levies a wealth tax, these records contain detailed wealth information for every member of the population. Our data are annual and span 26 years,

Figure 1.1: 2018 participation rate in Norway



Notes. This figure looks at the proportion of the population in 2018 who are currently in the stock market ("Participating") or not, where the latter are subdivided into those who have never participated before ("Never") and those who have participated at some point before ("Former"). Only those who appear in the data for at least 15 years are included in these calculations.

which is significantly longer than similar administrative datasets from other countries and of higher frequency than most relevant longitudinal surveys.[1] Individuals must file a tax return even if they hold no financial assets, which allows us to confidently identify periods of nonparticipation. This is a significant advantage relative to brokerage accounts data, where exit from such samples may simply reflect a transfer to another account provider rather than a complete withdrawal from the stock market. Financial holdings are directly reported to the tax authority by the financial intermediaries themselves. This third-party reporting alleviates concerns about measurement errors that can arise when using self-reported measures of wealth. In addition, we are able to link the tax records to other administrative datasets, thereby giving us additional information about each citizen that is typically not available in survey or brokerage accounts datasets.

We first document novel facts on two margins of stock market participation that have received less attention in the existing literature, namely the exit and reentry margins.[2] Our focus is on participation through nonretirement investment accounts, as existing work has documented inertia in retirement accounts (Brunnermeier and Nagel, 2008). We find that many individuals have very short spells in the stock market; that is, they stay in the stock market for only 1–2 years and then completely liquidate their stock holdings. 15%

[1] For example, the Swedish microdata on wealth used by Calvet et al. (2007, 2009a,b) cover the period from 1999 to 2007, and PSID survey waves are biennial.

[2] While our focus is on the speed of exit, others have linked exit to age (Poterba and Samwick, 1997; Ameriks and Zeldes, 2004; Fagereng et al., 2017a), house purchases (Brandsaas, 2021), income shocks (Bonaparte et al., 2021), and portfolio characteristics (Calvet et al., 2009a).

of all spells end just one year after exit, and a further 8% end in 2 years. We show that this behavior is neither due to liquidity needs, such as house purchases or unemployment, nor involuntary participation coming from inheritances or employee stock options. Our finding implies that the high exit rates documented in other studies are driven by new investors who invest for only a short period.[3]

We then investigate whether the likelihood of a short spell, which we define as a spell that results in complete exit within 2 years, is correlated with certain characteristics. Low income and wealth individuals, as well as those without a college degree, are more likely to exit quickly. Given the correlations between these characteristics and financial literacy (Lusardi and Mitchell, 2011; Behrman et al., 2012), this finding implies that lower financial literacy is associated with an increased likelihood of having a short spell. Men are 20% more likely to exhibit such behavior compared to women, supporting existing evidence that men tend to trade excessively and display overconfidence (Barber and Odean, 2001). We also find that quick exits are significantly more likely among investors who enter into directly held stocks rather than mutual funds. At the aggregate level, the prevalence of short spells in mutual funds rose during the bursting of the dot-com bubble in 2001–2002 and the financial crisis in 2008, whereas short spells in direct stockholding are more common during stock market booms. This distinction by asset class links to existing evidence on the disposition effect in individual stocks (Shefrin and Statman, 1985; Odean, 1998) and performance sensitivity in mutual fund flows (Ippolito, 1992; Gruber, 1996; Frazzini and Lamont, 2008). We then study how the probability of exit evolves with time spent in the market. Using the methodology of Alvarez et al. (2021), we estimate the hazard function for exit from participation to be downward sloping and convex, which means that the longer one stays in the stock market, the less likely they are to withdraw completely from the market. Together with the short spells result, this finding indicates that participation status is particularly fragile in the initial years following entry.

Moving onto the reentry margin, many individuals who completely exit from the stock market subsequently return. Over 35% of exiters reenter the stock market within the following four years, and they typically return to the same asset class (mutual funds or direct stockholding) that they previously invested in. Most reentry occurs soon after exit, often just a year later. We find that high income and wealth individuals are more likely to reenter. We also estimate a downward-sloping and highly convex hazard function for reentry, implying negative duration dependence in reentry probabilities: The longer an

[3] Bonaparte et al. (2021) show that, on average, 7.3% (8.7%) of year t households enter into (exit from) nonretirement investment accounts in year $t+2$. See also Hurst et al. (1998) and Vissing-Jørgensen (2002).

individual has been away from the stock market, the less likely they are to return. After about a decade of nonparticipation, the likelihood of reentry is effectively zero.

We then consider the implications of our empirical findings for theories of stock market participation. In particular, we examine the conditions under which a workhorse life-cycle portfolio choice model à la Cocco et al. (2005) can produce short-term dynamics. In this model, agents can invest in two financial assets, one risky (stocks) and the other safe (bonds), and they receive an exogenous labor income in every period that is stochastic during working life but constant in retirement. Under the core Cocco et al. (2005) model, there is full participation at all ages and thus no entry or exit dynamics.[4]

To generate a motive for nonparticipation, we augment the baseline model with per-period participation costs, which are a popular explanation for limited participation in the stock market (Vissing-Jørgensen (2002, 2003)). These costs are paid in every period in which one holds stocks and represent the opportunity cost of time needed to monitor and rebalance one's investment portfolio every period or, alternatively, broker management fees. In principle, per-period costs could generate both exit and reentry. If they are high relative to wealth, it may be optimal to fully liquidate when faced with adverse income or return shocks. Upon building up enough wealth, those who have exited may reenter. We find that the model requires sizable per-period participation costs of approximately $1,300 per annum to generate the degree of short-term dynamics observed in the data. This value is large relative to average holdings of public equity observed in the Norwegian data and is considerably higher than typical brokerage management fees charged in reality, as well as structural estimates of such costs by Fagereng et al. (2017a), Bonaparte et al. (2021), and Catherine (2021). High costs are required because precautionary savings motives are strong under reasonable degrees of risk aversion, leading to quick wealth accumulation that makes small costs redundant. Adding entry costs that capture the time and effort spent searching for an account provider or learning fundamental investment principles makes it even harder to match the data. This is because entry costs make exiting less attractive, given that these costs need to be repaid upon reentry. Consequently, adding entry costs requires an even higher per-period cost to generate sufficient exit and reentry.

To rationalize our facts under more plausible levels of participation costs, we extend the model to allow individuals to form beliefs over the equity premium based on realized

[4]Full participation is in line with the predictions of standard portfolio theory, which states that as long as the expected equity premium is positive, everyone should invest at least a small amount in stocks (Samuelson, 1969; Merton (1969, 1971)). This occurs because an individual with preferences exhibiting second-order risk aversion (e.g., CRRA utility) is essentially risk neutral with respect to small risks. As such, zero stockholding will be suboptimal given that the average equity premium is positive.

returns. This ingredient is motivated by the literature on memory and experience effects documenting how past experiences can have long-lasting effects on beliefs and actions (e.g., Greenwood and Nagel, 2009; Malmendier and Nagel (2011, 2015); Afrouzi et al., 2020; Bordalo et al., 2020).[5] In the model, agents lower their expectations of future stock returns upon receiving a poor return realization, making continued participation less attractive. We allow for a small degree of noise in belief formation to capture the impact of external signals coming from peers (Hong et al., 2004; Kaustia and Knüpfer, 2012), imperfect memory retrieval (Azeredo da Silveira et al., 2020), or cognitive limitations when forming beliefs (Fehr and Rangel, 2011).

The model with beliefs can explain the patterns of exit observed in the Norwegian data under much lower levels of participation costs. Short spells occur after poor initial returns, leading individuals to think that the return on stocks will be lower in the future. Some people exit because their expected equity premium becomes negative. If per-period participation costs are present, people may exit even if they perceive the equity premium to be positive on average because the expected return becomes too low relative to the costs. The inclusion of beliefs can also generate a downward-sloping hazard function for exit from participation. Individuals who have continuously participated for many years are likely to have experienced good returns, which is why they have not exited yet. As a result, they would require an extremely poor return to dampen their expected returns by a sufficient margin such that they wish to withdraw from the market. In line with the model, our data indicate that investors who exit soon after entry, on average, perform worse than longer-term participants; that is, they are more likely to report losses and less likely to report taxable gains.

The model can also rationalize the fact that many individuals have multiple spells. Reentry occurs in the model for two reasons. First, as we allow for small exogenous fluctuations in beliefs over time, some nonparticipants may reenter following a positive shock to their beliefs. For those with moderate beliefs, receiving such a nudge does not take long, resulting in quick reentry. However, people who are very pessimistic remain permanently out of the stock market. Second, in the presence of per-period participation costs, some individuals exit because their wealth is insufficient to justify paying the costs. However, upon accumulating further wealth, they may return. Our theory can also produce negative duration dependence in reentry probabilities because those who remain nonparticipants after many years following exit are likely to be individuals who performed so poorly in

[5]See Malmendier and Wachter (2021) for an overview of the empirical and theoretical literature on how experiences and memory affect choices.

their past spell that their expected return on equities is too weak to warrant reentry. We provide supporting evidence for this behavior by showing that prior losses are more common among reentrants than among those who chose not to return.[6]

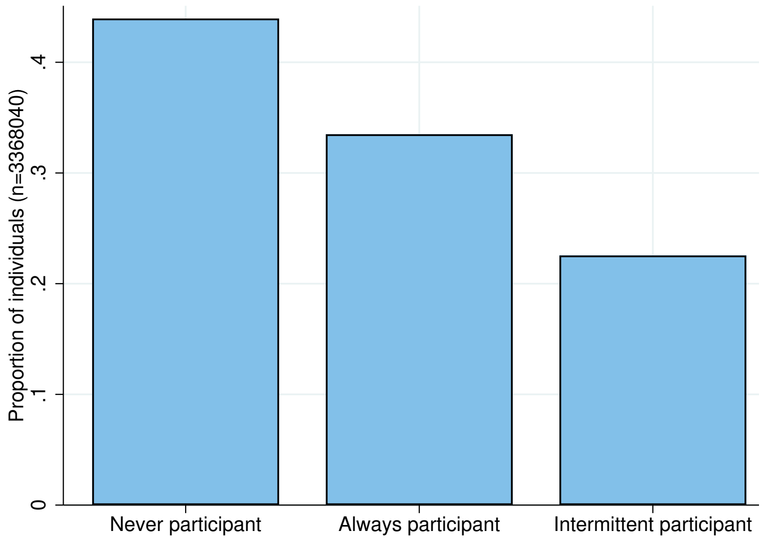
Our findings contribute to the broad literature on underparticipation in the stock market by retail investors (Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995; Vissing-Jørgensen (2002, 2003); Campbell, 2006; Choi and Robertson, 2020). We approach this puzzle from a dynamic perspective. While the literature typically divides the population into two groups, namely those who never invest in the stock market and those who continually invest, we find that many individuals fall into a third category of being *intermittent* participants. At least 20% of the Norwegian population belongs to this group, which consists of people who either participate once for only a short period (< 5 years) or have multiple spells (Figure 1.2). Therefore, a snapshot of an individual's participation decision in a single year is not necessarily representative of their choices at other points. Furthermore, we are also able to shed light on existing theories of participation used to rationalize the underparticipation puzzle.[7] We show that a necessary condition for the workhorse portfolio choice model of Cocco et al. (2005) augmented with fixed participation costs to generate the patterns we observe is that per-period costs must be very high.

Our results can have important implications for wealth accumulation. A growing literature has established a link between portfolio choices and wealth inequality (Benhabib et al., 2011; Gabaix et al., 2016; Benhabib et al., 2019; Bach et al., 2020; Hubmer et al., 2021; Xavier, 2021). We find that many individuals have intermittent spells in the stock market, so they may not remain in the stock market for long enough to attain the large average equity premium. Although efforts have been made to boost stock market participation (e.g., through tax incentives), our findings indicate that it is not simply about encouraging entry. Perhaps individuals need to be encouraged to continue participating for a prolonged period, particularly when faced with poor short-term returns. Furthermore, we find that individuals of low wealth are more likely than wealthier individuals to exit the stock market soon after entry, which can further exacerbate the wealth gap between these two groups.

[6]The lower propensity to return following a bad return in the model relates to empirical evidence in Kaustia and Knüpfer (2008) and Chiang et al. (2011), who find that personally-experienced returns on past IPOs can affect the likelihood of participating in subsequent auctions.

[7]Gomes et al. (2021) groups explanations for the underparticipation puzzle into four categories, namely participation costs (Haliassos and Bertaut, 1995; Haliassos and Michaelides, 2003; Vissing-Jørgensen (2002, 2003); Choi and Robertson, 2020), nonstandard preferences (Epstein and Zin, 1990; Epstein and Wang, 1994; Segal and Spivak, 1990), risks faced by households (Benzoni et al., 2007), and cultural/social factors (Hong et al., 2004; Kaustia and Knüpfer, 2012).

FIGURE 1.2: Types of individuals



Notes. This figure divides individuals into three categories based on their lifetime stock-market exposure and plots the percentage of the population belonging to each of these three groups. "Never participant" contains individuals who never hold any stocks. "Always participant" covers people who are observed to have one single spell lasting at least 5 years, plus any individuals with right-censored or left-censored spells. "Intermittent participant" contains individuals with one single spell known to last less than 5 years and people who have multiple spells in the stock market. This figure includes only individuals observed in the data for at least 15 years.

Outline: The paper is structured as follows. Section 2 describes the Norwegian data, while Section 3 documents our exit and reentry facts. Section 4 details the workhorse portfolio choice model and our augmented model with beliefs. Section 5 shows the model results, and Section 6 concludes.

2 Data

We use Norwegian administrative data to conduct our analysis. Most administrative datasets contain information on income only. However, due to the existence of a wealth tax in Norway, our data also contain detailed information on wealth holdings by broad asset class for each resident as of December 31 for each year from 1993 to 2018. The Norwegian data are particularly well suited to studying individual-level dynamics in stock market participation relative to other datasets. First, to study dynamics, we need to be able to follow individuals over time. Compared to other datasets, our data provide this panel dimension with a longer time dimension.[8] Second, a concern with brokerage accounts data is that exit from the sample does not necessarily mean an exit from the stock market. For example, if an individual simply switches providers, they would appear as an

[8] As the wealth tax in Sweden ended in 2007, the Swedish data used by Calvet et al. (2007, 2009a,b) span 9 years (1999–2007 inclusive). The brokerage data of Barber and Odean (2000, 2001) cover 1991–1996.

exiter in the brokerage accounts data. Reentrants could be difficult to identify if account numbers change between spells. The Norwegian data do not have this concern, as the tax data are based on overall holdings across all financial intermediaries and identification is at the individual level. Third, brokerage data can have concerns with sample selection and nonrandom attrition, the latter of which is also a worry with longitudinal survey data. The Norwegian data cover the full population, and attrition should be due to death or emigration only. Fourth, financial institutions directly report information on wealth holdings to the tax authority, which eases concerns about measurement errors.[9] As such, evading taxes in Norway by underreporting asset holdings is very challenging.[10] Last, we are able to link the tax records to other administrative datasets, which contain additional information about each individual that is not necessarily available in survey or brokerage accounts data (e.g., demographics, employment history, and house purchases). This allows us to study whether the behaviors we find are linked to certain characteristics.

While the Norwegian data are particularly promising for our research objective, they have their shortcomings. The data provide us with asset holdings as of December 31 of each year. As such, we are limited to participation decisions at the annual frequency, although it is worth noting that this is more frequent than most panel survey waves.[11] We are, therefore, unable to capture within-year spells, although the presence of within-year spells would strengthen our result that short spells in the stock market exist. In addition, we do not have information on occupational or public pension wealth. However, in Section 3.3.3, we argue that pensions are unlikely to affect our results. Third, we do not have information on the specific mutual funds held.

2.1 Data construction

We use the tax records to construct wealth by broad asset class and combine them to obtain measures of financial and real wealth.[12] Financial wealth can be decomposed into the following asset classes: (a) cash and deposits (both domestic and foreign), (b) directly held listed stocks, (c) directly held unlisted stocks (typically private equity), (d)

[9]Following this direct reporting, residents are sent a prefilled tax form to approve. If they do not respond, then the tax authority assumes that the information is correct. In 2009, around 60% of tax payers in 2009 did not respond (Fagereng et al., 2017a).

[10]As noted by Fagereng et al. (2017a), one source of under-reporting would be if individuals hold but fail to disclose foreign investments. While asset holdings through Norwegian financial intermediaries are directly reported, this is not the case for foreign holdings. For Sweden, Calvet et al. (2007) argue that such holdings are likely to be a small portion of overall assets other than for the wealthiest individuals.

[11]For example, wealth information in the PSID was collected from 1984 at 5-year intervals until 1999, when it began to be collected biennially.

[12]We provide further details on the construction of the wealth variables in Appendix A.

stock mutual funds, (e) money market funds, (f) financial wealth held abroad, and (g) other financial assets.[13] Real wealth consists of housing and other real assets.[14] We are most interested in the extensive margin of participation and treat an individual to be participating in a given year if any of directly held listed stock holdings, stock mutual fund holdings, or financial wealth held abroad are strictly positive.[15] We focus on stock market participation through nonretirement investment accounts because there is typically little turnover and trading activity in retirement accounts (Brunnermeier and Nagel, 2008; Bonaparte et al., 2021). We also restrict attention to individuals aged 20 or over to ensure that the person is the main asset holder.[16]

2.2 Descriptive statistics

Figure 1.3 plots the stock market participation rate in Norway over time. Only 12% of the population owned stocks at the start of the sample. However, there was an acceleration in participation during the 1990s. Reasons include improved access to financial markets for retail investors, the rise of mutual funds, and the popularity of technology stocks during the dot-com bubble (Guiso et al., 2003a).[17] After the bursting of the dot-com bubble, the participation rate dropped sharply. It stabilized until the financial crisis and has since shown a persistent decline. Figure 1.4 plots the entry and exit rates over time. The sharp fall in participation in the early 2000s can be linked to a pronounced rise in the exit rate and a drop in the entry rate. Since the financial crisis, entry rates have been particularly low, which can explain the downward trend in the participation rate.

Table 1.1 provides summary statistics at the individual level. The first block shows that there is an even split of men and women in the sample, and 36% of individuals have a college degree. The second block provides information on income and wealth holdings. The average individual has a total gross wealth holding of $277,000, though the large standard deviation in asset holdings illustrates the vast heterogeneity in wealth across the population. The median wealth holding is less than half of the mean holding, indicating

[13] Other financial assets consist of outstanding claims and receivables, shares of capital in housing cooperatives or jointly owned property, own pension insurance and life insurance, and other wealth.

[14] Other real assets include vehicles (e.g., boats, cars, caravans), holiday homes, fixtures and other business assets, contents, and other real estate (e.g., farms, plots).

[15] We include financial wealth held abroad in this definition to be conservative because the nature of such wealth is not observed, and hence it could contain foreign stockholdings. However, few people report holding wealth abroad ($<$ 2% of observations).

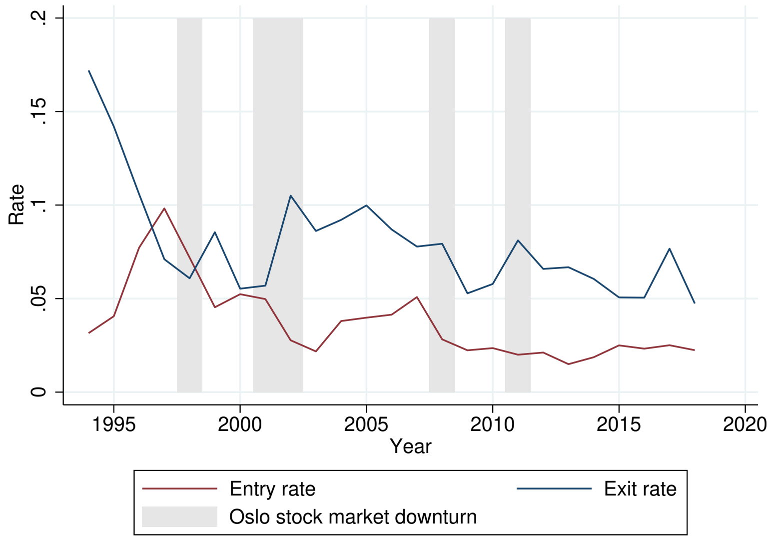
[16] Fagereng et al. (2020) also impose an upper bound on age of 75 in their study of return heterogeneity. However, we do not do so, as it can artificially generate right-censored spells.

[17] Figure 1.22 shows the participation rates separately for stock mutual funds and directly held stocks. Participation in mutual funds rose by more than fivefold from 1993 to the early 2000s. Participation in directly held stocks also rose but by a smaller margin: from just over 8% in 1993 to around 12% in 2000.

Figure 1.3: Stock market participation rate over time

Notes. This figure plots the participation rate in the stock market annually from 1993 to 2018.

Figure 1.4: Entry and exit rates over time



Notes. This figure plots the entry and exit rates for stock market participation. The entry rate in year t is the proportion of nonparticipants in year $t-1$ who enter in year t. The exit rate in year t is the proportion of participants in year $t-1$ who leave the stock market in year t. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

a rightward skew in the wealth distribution. Nonfinancial wealth, of which the major component is housing, accounts for a larger share of total wealth than financial wealth does, with the average individual holding \$80,000 in financial wealth compared to \$196,000 in nonfinancial wealth. The mean amount of wealth held in public equity, measured as the sum of holdings in stock mutual funds, directly held stocks, and financial wealth abroad, is just over \$7,000. Indeed, the median individual does not hold any public equity, a finding that is indicative of broad aggregate underparticipation in the stock market in Norway. The third block further verifies this finding by showing that 25% of individuals invested in the stock market in 2018. Most participants invest in mutual funds rather than directly holding stocks. Conditional on participating in the stock market, 26% of financial wealth is in stocks on average.

TABLE 1.1: Summary statistics

	Mean	Std. dev	P10	Median	P90	P99
Demographics						
Age (in years)	48.80	18.54	26	47	74	91
Male	0.52	0.50	0	1	1	1
Single	0.34	0.47	0	0	1	1
College degree	0.36	0.48	0	0	1	1
Income and wealth ($'000s)						
Gross income	40.66	48.76	0	36.81	88.66	177.71
Financial wealth	80.04	1,984.23	0.01	10.60	123.79	835.37
Financial wealth in public equity	7.05	137.09	0.00	0.00	8.42	122.72
Nonfinancial wealth	196.49	299.93	0	131.11	487.62	1,185.16
Gross wealth	276.53	2,063.84	0.03	165.27	593.29	1,791.39
Net wealth	185.92	2,025.55	-38.51	42.54	479.71	1,565.64
Participation and wealth shares						
Participates in public equity	0.25	0.43	0	0	1	1
Participates in mutual funds	0.22	0.41	0	0	1	1
Participates in individual stocks	0.06	0.24	0	0	0	1
Cond. risky share (% of total wealth)	8.32	15.64	0.10	2.26	23.23	84.00
Cond. risky share (% of fin. wealth)	26.29	26.77	1.01	16.22	70.32	97.26
Observations	4.64m					

Notes. This table provides summary statistics using data from 2018. The first block gives summary statistics for demographic characteristics. *Single* is a binary variable equal to 1 if the individual is neither married nor cohabiting, and zero otherwise. The second block information on income and wealth measured in USD (in thousands) based on an exchange rate of $1=8.64 NOK at the end of 2018. *Gross income* is income from all sources. *Public equity* is measured as the sum of holdings in stock mutual funds, directly held stocks and financial wealth abroad. The third block gives summary statistics on stock market (i.e., public equity) participation and the share of wealth invested in public equity conditional on holding a nonzero amount of such wealth.

3 Empirical facts

In this section, we study two margins of stock market participation using the Norwegian administrative data, namely exit (Section 3.1) and reentry (Section 3.2). In Section 3.3, we discuss and rule out potential explanations of our findings.

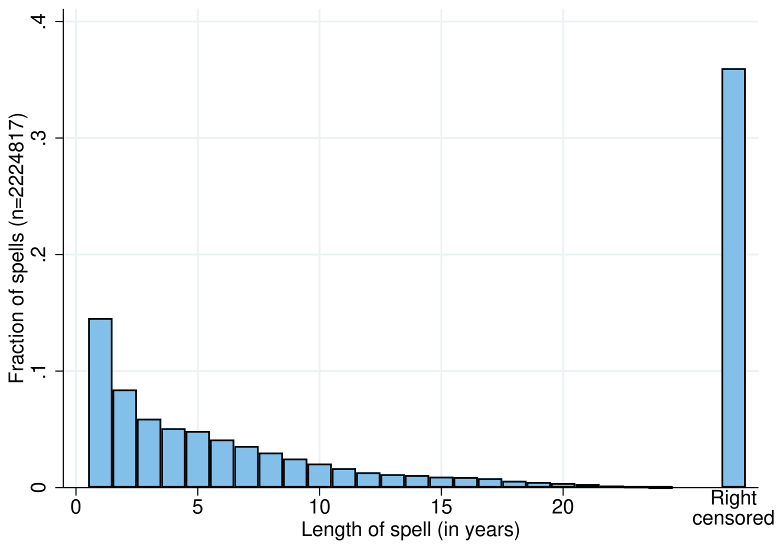
3.1 Exit margin

3.1.1 Short spells are common

We begin by examining the distribution of spell lengths in the data. Figure 1.5 plots a histogram with the distribution of spell lengths based on spells beginning between 1994 and

2015 inclusive.[18] We restrict attention to spells starting by 2015 to ensure that participants have at least 3 years in which to exit. If, for example, 2017 entrants were also included, they would either have a 1-year completed spell or be right censored. Thus, including such entrants would artificially inflate the bars corresponding to a short spell length. The histogram shows a declining relationship between spell length and the proportion of observations. Almost 15% of all spells end in just 1 year, and 23% end within 2 years. We undertake a variety of robustness checks, namely analysis at the household level (Figure 1.23), excluding entrants who receive a gift or inheritance in the year of or before entry (Figure 1.24), removing individuals with stocks in the company they work for (Figure 1.25), dropping investors who invest a small sum at the point of entry (Figure 1.26), and only using the first (recorded) spell for each participant (Figure 1.27). In all cases, similar patterns emerge.

FIGURE 1.5: Distribution of spell lengths



Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data. We take all spells beginning at any point from 1994 to 2015. The x-axis gives the spell length (in years), and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

The next step is understanding whether short spells can be linked to observable characteristics. To do this, we estimate the following linear probability model:

$$\Pr(\text{spell ends within 2 years}) = \alpha_i + \delta_t + \beta' X_{it} + \epsilon_{it} \tag{1.1}$$

where δ_t denotes entry-year fixed effects and X_{it} is a vector of observable characteristics measured at the point of entry, such as age and wealth. Given that we observe individuals

[18]Left-censored spells are excluded from this figure, as a spell length cannot be computed for such spells. These spells are typically those that were already ongoing at the start of our data, though other reasons for left-censoring could be immigration of an existing stockholder into Norway.

with multiple spells, we are able to include individual fixed effects α_i to absorb time-invariant characteristics.

TABLE 1.2: Determinants of short spells ($\leq$2 years)

Male	0.047***	
	(0.001)	
College degree	-0.008***	-0.023***
	(0.001)	(0.004)
Homeowner	0.001	-0.009**
	(0.001)	(0.003)
Unemployed	0.017***	0.000
	(0.001)	(0.003)
Single	0.023***	0.016***
	(0.001)	(0.002)
Directly held stocks	0.083***	0.091***
	(0.001)	(0.002)
Sample mean	0.23	0.36
Individual FE	No	Yes
Entry year FE	Yes	Yes
Age group FE	Yes	Yes
Income decile FE	Yes	Yes
Wealth decile FE	Yes	Yes
Observations	2242427	866406
R-squared	0.04	0.47

Notes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. This table shows the results from estimation of Equation 1.1. The first column excludes individual fixed effects, while the second column includes them. The dependent variable is a binary variable equal to 1 if the spell ends within 2 years, and zero otherwise. *Homeowner* is a binary variable equal to 1 if the participant owns their own property (either self-owned or ownership through housing cooperatives), and zero otherwise. *Single* equals 1 if the participant is neither married nor cohabiting, and zero otherwise. *Unemployed* equals 1 if the participant receives unemployment benefits at the point of entry, and zero otherwise. *Directly held stocks* equals 1 if the participant buys stocks directly at the point of entry, and zero otherwise. Entry-year fixed effects are included. Age fixed effects by broad age group (20-29, 30-39, 40-49, 50-59, 60-69 and 70+), as well as income and wealth decile fixed effects are included. Observables are measured at the point of entry. Standard errors are clustered at the individual level. The regression uses data on entrants from 1994-2016.

Table 1.2 shows the result from this estimation in specifications with and without individual fixed effects. Men are 20% (4.7pps) more likely than women to have a short spell. This result relates to the existing literature on gender differences in confidence and trading behavior, which has found that men tend to be more overconfident and trade excessively, often to the detriment of their own returns (Barber and Odean, 2001).[19] We also see that single individuals are more likely to have short spells, in line with the view that married individuals can influence the investment decisions of their spouse and thus reduce the effect of individual overconfidence (Barber and Odean, 2001). In addition, entrants who invest in stocks directly are 40% (9.1pps) more likely to have a short spell relative to those who invest in mutual funds. Characteristics typically associated with lower financial literacy

[19] Grinblatt and Keloharju (2009) study overconfidence using Finnish data and show that individuals with a high degree of self-confidence tend to have higher trading volumes.

are also linked to a higher prevalence of short spells.[20] Having a college degree lowers the likelihood of exiting within 2 years by 10% (2.3pps). Figures 1.6a and 1.6b show the impacts of income and wealth, respectively.[21] For income, we find a monotonic negative relationship between income and the probability of a short spell, with those in the bottom income decile having an 11% (2.5pps) higher probability of a short spell relative to the median. For wealth, the impact of low wealth is even more striking. Entrants belonging to the bottom wealth decile are 43% (10pps) more likely to exit within 2 years relative to the median.[22] For wealth levels above the median, there is no significant difference in the prevalence of short spells, indicating that it is low wealth that particularly matters. Calvet et al. (2009a) show that individuals with less income, wealth, and education are more likely to exit. Our finding suggests that they are not just more likely to exit at any point. Rather, they are also more likely to experience a quick exit. Taken together, short spells are more prevalent for individuals with characteristics linked to lower financial literacy. Regarding age, we see that short spells are more likely for the youngest and oldest age groups (Figure 1.29). This finding is in line with Fagereng et al. (2017a), who show that younger households tend to enter and exit frequently. It is important to note, however, that short spells are not exclusive to these subgroups. Indeed, Figure 1.30 shows the distribution of spell lengths by income, wealth, education, gender, and asset class. For example, while men are more likely to exit quickly (Figure 1.30d), over 20% of women still leave the stock market within 2 years of entry. Thus, short spells are widespread and not purely concentrated among a particular subpopulation, though they are *relatively* more likely for certain groups. Figure 1.32 plots the average risky share at the point of entry for individuals of different eventual spell lengths, and shows that short spellers invest slightly more as a share of their wealth at entry relative to longer-term participants.

Our finding that short spells in the stock market are common can have important implications for wealth accumulation. Indeed, much of the policy focus has been on encouraging entry into the stock market (e.g., via tax incentives). However, we see that temporary participation is very common, so from a policy perspective, it is not only about encouraging entry into the stock market. It is also important to encourage participants not to exit impulsively to ensure that they earn the high equity premium on average.

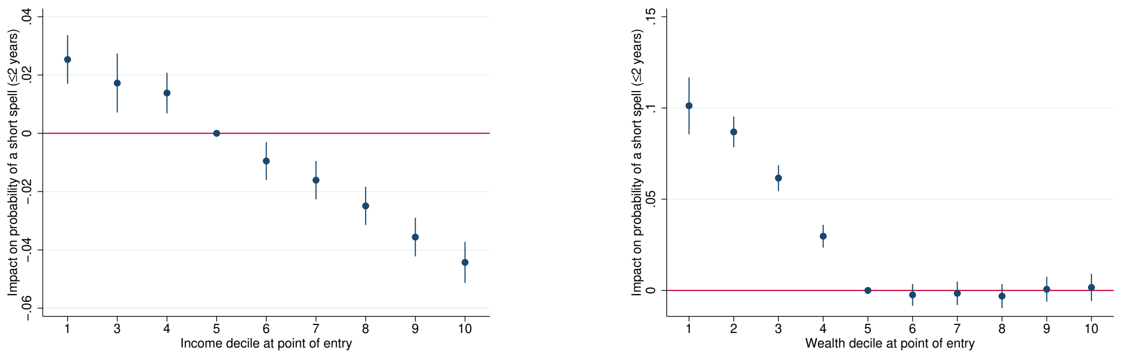
[20] Lusardi and Mitchell (2011) give evidence of a positive correlation between educational attainment and financial literacy. Behrman et al. (2012) find this as well and further show a positive correlation between wealth and financial literacy.

[21] In Figure 1.6a, there is no 2[nd] decile for income because > 20% of observations have zero income, and these are all grouped in the first decile. The first decile can, therefore, be thought of as a zero-income group. This will also be the case in subsequent plots involving income.

[22] Figure 1.28 plots the wealth distribution for short and longer-term spellers separately, and further shows that short spellers are more likely to belong to a lower wealth decile than longer-term participants.

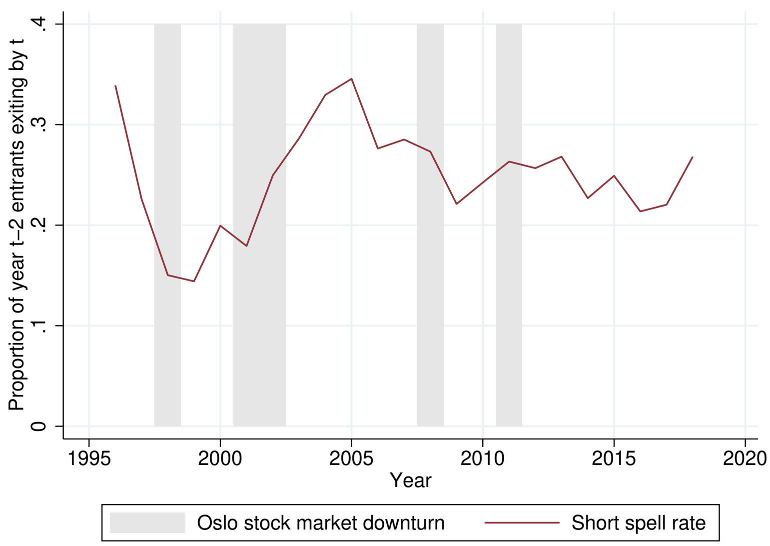
Figure 1.6: Impact of income and wealth on the probability of a short spell

(A) Income (B) Wealth



Notes. This figure plots the coefficient estimates for the fixed effects on income (A) and wealth (B) deciles following the estimation of Equation 1.1 with individual fixed effects. Variables are measured at the point of entry, and deciles are based on the full Norwegian population aged 20 and above in that year. The effects are estimated relative to the median group. 95% confidence intervals are shown. The red line represents a null relative effect.

Figure 1.7: Prevalence of short spells over time



Notes. This figure plots the proportion of year $t-2$ entrants who leave the stock market by year t. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

Figure 1.7 plots the prevalence of short spells over time. We see that short spells became much more likely during the early 2000s, a period that coincides with the booming stock market and subsequent bursting of the dot-com bubble. Indeed, this period exhibited significant trading volumes and stock market inflows and outflows (Ofek and Richardson, 2003; Hong and Stein, 2007). However, the timing of short spells differs based on the asset class. For mutual funds (Figure 1.31a), this prevalence of short spells rises during stock market downturns, such as the bursting of the dot-com bubble (2001–02), the financial crisis (2008), and the European sovereign debt crisis (2011). Instead, for direct stockholding (Figure 1.31b), short spells were most prevalent in the late 1990s and the mid-2000s, periods when the stock market was booming. This distinction relates to findings in Calvet et al. (2009a) that households are more likely to exit direct stockholding when their individual stocks have performed well in line with the disposition effect.[23] However, households are more likely to exit from mutual funds following poor returns.[24] Therefore, short spells in mutual funds are more likely during stock market downturns, whereas they are more prevalent during booms for direct stockholding.[25]

3.1.2 Exit probability falls with spell duration

Are investors more likely to exit the stock market in the initial periods following entry or after staying in the market for a prolonged period? To answer this question, we estimate the hazard function for exit from participation. The hazard function $h_i(d)$ denotes the probability that individual i exits the market d years after entry, conditional on not exiting until then. A standard challenge with hazard function estimation is separating true duration dependence from (unobserved) heterogeneity. Estimating hazard functions based on pooled samples with heterogeneous individuals can lead to a downward bias in the slope of the hazard function because individuals who are less likely to "survive" exit the sample earlier than others (Lancaster, 1979; Kiefer, 1988).

To address this concern, we apply the linear GMM estimator of Alvarez et al. (2021) and estimate a discrete-time proportional hazard model of duration. The main advantage of this approach is that it gives a consistent estimator of the slope of the hazard function, even in the presence of time-invariant individual heterogeneity. The Alvarez et al. (2021) methodology can do so by exploiting the presence of individuals with multiple spells in the stock market. The resulting limitation is that the set of people experiencing multiple spells used in the estimation can be fundamentally different from the rest of the population.[26] However, similar patterns do emerge when using the full set of participants and instead estimating a Cox proportional hazards model (Figure 1.33).[27] Further details on the

[23]The disposition effect refers to the tendency of investors to sell stocks trading at a gain and to hold stocks trading at a loss (Shefrin and Statman, 1985; Odean, 1998).

[24]The literature has found a strong positive correlation between mutual fund flows and past performance (Ippolito, 1992; Gruber, 1996; Frazzini and Lamont, 2008).

[25]Figure 1.38 plots the exit rates for the two asset classes separately and shows similar time-series patterns to the prevalence of short spells.

[26]Table 1.4 shows that individuals with multiple spells look more similar to "always participants" (individuals who have one single spell lasting at least 5 years) rather than participants with one short spell.

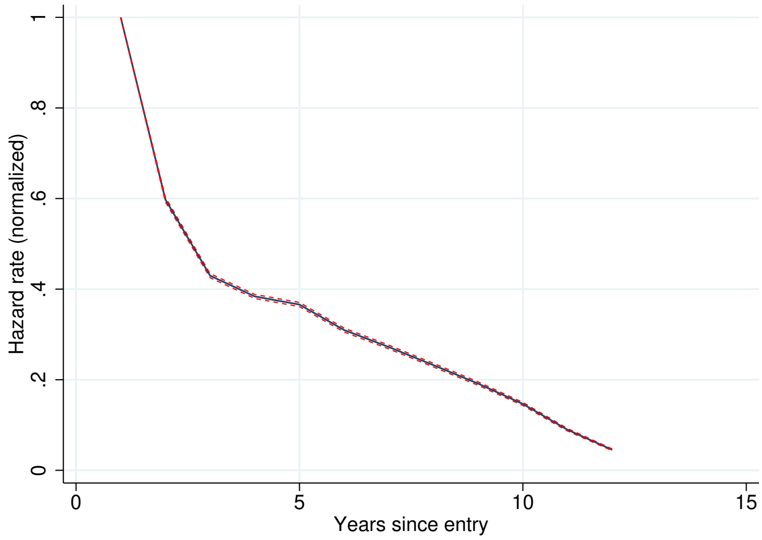
[27]The Cox proportional hazards model takes the form:

$$h_i(d) = \exp{(X_i\beta)}b_d \tag{1.2}$$

where X_i is a set of individual characteristics and b_d is the baseline hazard. Estimated hazard ratios (exponentiated coefficients) for the covariates X_i are given in Table 1.5 and Figures 1.34-1.36.

Alvarez et al. (2021) approach are provided in Section C.

FIGURE 1.8: Baseline hazard function for exit from participation



Notes. This figure plots the estimated baseline hazard for exit from participation following the methodology of Alvarez et al. (2021) described in Section C. The dotted red lines denote 95% confidence intervals. The hazard rate at duration $d = 1$ is normalized to 1.

Figure 1.8 plots the estimated baseline hazard function. The hazard function is monotonically declining in duration, indicating negative duration dependence; that is, the longer one has been participating in the stock market, the less likely they are to exit completely at that point in time. As described in Section C, we are able to recover the slope of the baseline hazard rather than its level using the Alvarez et al. (2021) approach, so we normalize the hazard rate at $d = 1$ to 1. A striking feature of the hazard function is the steepness of the slope in the initial years following entry. The hazard rates at $d = 2$ and $d = 3$ are 60% and 40% that of $d = 1$, respectively. By $d = 12$, the hazard rate is close to zero, suggesting that if an individual remains in the market for a prolonged period, the likelihood of them completely withdrawing from the market is minimal. Combined with the fact that many stock market participants stay in the stock market for a short time, this finding indicates strong dynamics in the initial years following entry.

3.2 Reentry margin

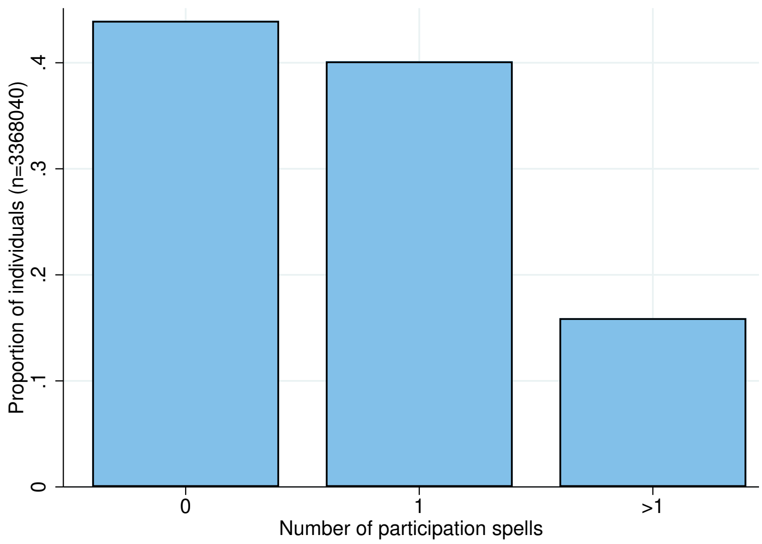
3.2.1 Many exiters reenter the stock market

We now turn to understanding whether exiters ever reenter following exit. Figure 1.9 plots the distribution of the number of spells an individual experiences.[28] 43% of the

[28] We restrict attention to individuals who appear in the data for at least 15 years, as those who appear for fewer years are likely to have either zero or one spell, which would skew the distribution to the left.

population never participates in stocks, while 40% are observed to participate just once, meaning that around 17% of the entire population has multiple spells. Hence, reentry does occur for a sizeable fraction of people. Indeed, this finding negates the conventional view in the literature that upon entering the stock market, individuals should rarely leave. Here, we see that a large share of people liquidate their stock holdings completely but subsequently reenter.[29] We also find that investors tend to return to the same asset class in which they previously participated. Figure 1.39 shows that over 80% of reentrants who previously participated in funds choose to return to funds. Of those reentrants who previously invested in individual stocks, over 60% go back into direct stockholding. This result suggests that investors tend to divide themselves into types, namely fund investors and direct stockholders, with very few participating in both.

FIGURE 1.9: Number of spells



Notes. This figure plots the distribution of the number of spells using individuals who appear in the data for at least 15 years.

We now examine which characteristics are associated with reentry. For this purpose, we run the following linear probability model:

$$\Pr(\text{reenter within 4 years}) = \alpha_i + \delta_t + \beta' X_{it} + \epsilon_{it} \tag{1.3}$$

where δ_t now denotes exit-year fixed effects. We use a fixed window of 4 years to reenter because those who exit early in the sample have more years remaining in which to reenter. A fixed window means all exiters have the same amount of time to reenter. Furthermore, to preview the findings in Section 3.2.2, most reentry occurs soon after exit, and so a 4-year window should capture a large proportion of reentry. To ensure that all individuals

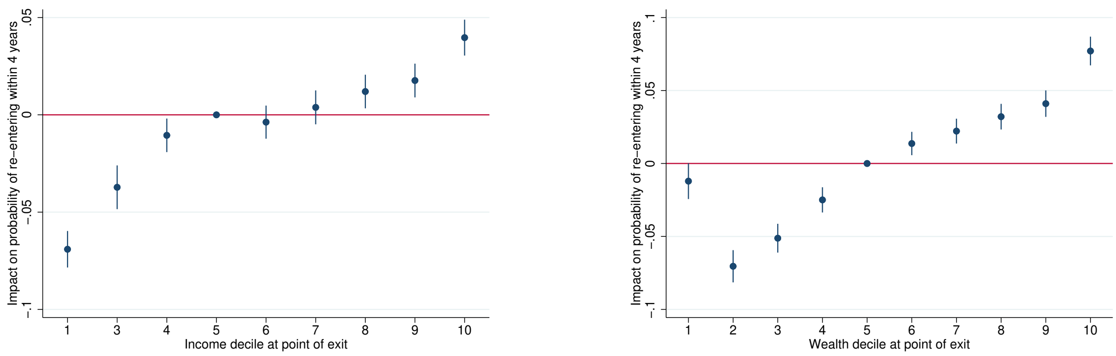
[29]Figure 1.40 decomposes the entry rate into reentrants and new entrants, and shows that about one-third to one-half of entrants in a given year are reentrants.

are observed for at least 4 years following exit, we restrict attention to those who left the stock market by 2014.[30]

Figure 1.10 plots the estimated effects of income and wealth and shows that reentry is more likely for individuals with higher income and wealth. The top income decile is 11% (4pps) more likely to reenter relative to the median (Figure 1.10a), and the highest wealth decile group is about 23% (8pps) more likely to reenter relative to the median (Figure 1.10b). Indeed, Calvet et al. (2009a) show that entry is more likely for individuals with high income and wealth.[31] Reentry is less likely for the youngest and oldest age groups (Figure 1.37), in line with the finding in Fagereng et al. (2017a) that permanent exit rises sharply after retirement. The estimated coefficients for the other variables are provided in Table 1.6.

FIGURE 1.10: Impact of income and wealth on the probability of reentry

(A) Income (B) Wealth



Notes. This figure plots the coefficient estimates for the fixed effects on income (A) and wealth (B) deciles following the estimation of Equation 1.3 with individual fixed effects. Variables are measured at the point of exit, and deciles are based on the full Norwegian population aged 20 and above in that year. The effects are estimated relative to the median group. 95% confidence intervals are shown. The red line represents a null relative effect.

3.2.2 Reentry often occurs soon after exit

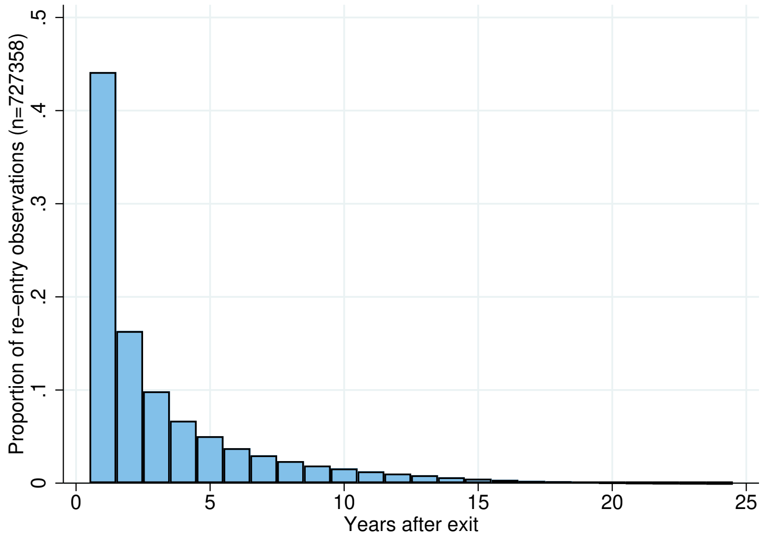
Conditional on occurring, how soon after exit do individuals reenter? Figure 1.11 plots the distribution of reentry times observed in the data. Almost half of all reentry occurs just 1 year after exit, indicating that reentry tends to be quick. Combined with the evidence for short spells given in Section 3.1.1, this implies that there is a high degree of turnover between participation and nonparticipation states, with many individuals dropping out of

[30]Figure 1.41 plots a time series of the reentry rate. It is highest at the start of our sample but remains fairly steady from around 1998 onward, though there are drops during the stock market crashes in 2001–02 and 2008.

[31]They also find a positive effect of education. We find a similar positive effect of a college education in the specification without individual fixed effects, but the coefficient becomes insignificant after including them.

participation spells after only a few years and a nonnegligible number reentering soon after exit. These findings are robust to excluding recipients of gifts or inheritances (Figure 1.42) and individuals holding stocks in the company they work for when they reenter (Figure 1.43).

FIGURE 1.11: Distribution of reentry times



Notes. This histogram shows the distribution of reentry times in the Norwegian data.

3.2.3 Probability of reentry falls with time since exit

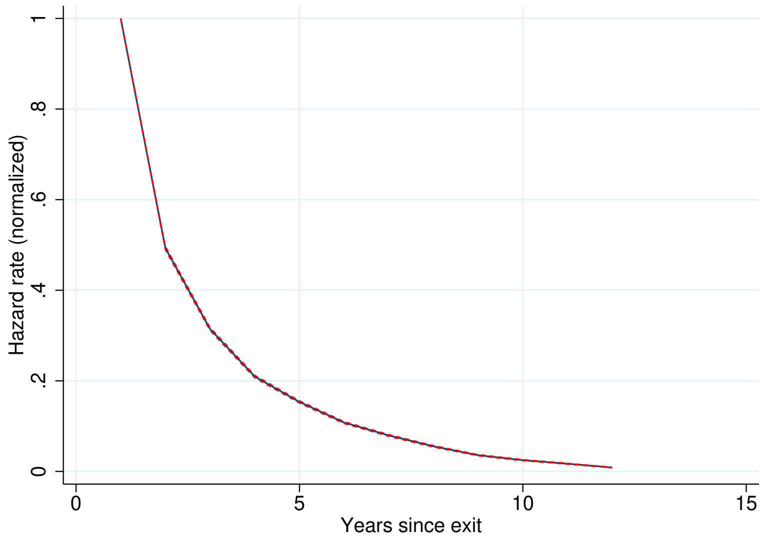
Our final fact studies how the likelihood of reentry changes with the duration since exit. Our object of interest is the hazard function $h(d)$, which denotes the probability of reentering d years after exit conditional on not having reentered until then. We exploit the fact that some individuals have multiple spells out of the stock market and again apply the GMM estimator of Alvarez et al. (2021). Figure 1.12 plots the estimated hazard function for reentry. The hazard function is downward sloping and highly convex, indicating negative duration dependence in reentry following exit: The longer it has been since one has been out of the stock market, the less likely they are to return. There is a sharp decline in the hazard rate in the initial years following exit, with the hazard rate at $d = 2$ being less than half that of $d = 1$. By $d = 12$, the hazard rate is very low, indicating that the likelihood of reentering is virtually zero by about a decade after exit.[32]

[32] Similar patterns appear if we estimate a Cox proportional hazards model for reentry (Figure 1.44), which takes the form:

$$h_i(d) = \exp(X_i\beta)b_d \tag{1.4}$$

where X_i is a set of individual characteristics and b_d is the baseline hazard. Estimated hazard ratios (exponentiated coefficients) for the covariates X_i are given in Table 1.7 and Figures 1.45-1.47.

FIGURE 1.12: Baseline hazard function for reentry



Notes. This figure plots the estimated baseline hazard for reentry following exit using the methodology of Alvarez et al. (2021) described in Section C. The dotted red lines denote 95% confidence intervals. The hazard rate at duration $d = 1$ is normalized to 1.

3.3 Ruling out potential explanations

In this section, we assess whether the quick exit and reentry patterns that we observe could be explained by certain factors, namely liquidity shocks, sophisticated market timing, pensions, or tax optimization.

3.3.1 Liquidity shocks

In principle, individuals might have to leave the stock market due to liquidity needs. For example, people may lose their job or face unexpected health expenses. Upon the "completion" of such liquidity needs, individuals may subsequently reenter the market. In general, one might expect a constant Poisson arrival of such shocks. However, a constant arrival rate would imply a flat hazard of exit from participation, which contradicts the downward-sloping hazard estimated in Figure 1.8. One might even expect liquidity needs to generate an upward-sloping hazard, as people would likely not enter the stock market in the year before any expected liquidity needs (e.g., house purchase) given the risk of a stock market downturn. To directly verify that liquidity needs are unlikely to drive our results, we undertake two checks. First, we identify observable liquidity needs in the data and study whether their prevalence varies with spell length. In particular, we look at house purchases, divorce, unemployment, and a large drop in income ($> 50\%$) as our liquidity shocks.[33] Figure 1.49 plots the proportion of exiters of different spell lengths experiencing at least one of these shocks in their exit year. For comparison, we also show the proportion of continuing participants experiencing a liquidity shock. We can see that some exit is correlated with such needs: about 10% of nonexiters experience a liquidity shock compared to around 15% for exiters. This is in line with papers linking exit to house purchases (Brandsaas, 2021), marital status (Christiansen et al., 2015), and unemployment (Basten et al., 2016). However, the prevalence of liquidity shocks is very similar across spell lengths, suggesting that short spellers do not have a higher likelihood of facing a liquidity shock compared to longer spellers. Furthermore, if 15% of exiters leave because of one of these observed shocks, it means that 85% of exiters are leaving for other reasons.

Second, we examine what happens to other components of the balance sheet at the point of exit. If individuals face liquidity needs, then presumably they would withdraw from their safe liquid asset holdings first. However, Figure 1.50 shows that the growth rate of safe financial asset holdings is actually much higher in the exit year relative to other years.[34] We see little movement in consumption growth (Figures 1.52-1.53) or nonhousing real assets (Figures 1.56-1.57), and only a small increase in the probability of house purchases (Figures 1.54-1.55) in the year of exit. This suggests that many exiters tend to put the funds into their bank account when they leave the stock market. All together, it appears that liquidity shocks are unlikely to explain the prevalence of short and multiple spells in the stock market.

3.3.2 Sophisticated market timing

Could the short-lived entry and exit observed in the data be driven by sophisticated market timers? Perhaps these individuals pursue short-term investment strategies and reenter whenever a promising investment opportunity arises. If this were the case, we would expect short spelling to be correlated with proxies for financial sophistication. However, Table 1.2 and Figure 1.6 show that short spelling is negatively correlated with characteristics typically associated with higher financial literacy (college education, income and wealth). Furthermore, we might expect higher returns for more sophisticated investors. However,

[33]Two other liquidity needs could be health shocks and education costs. However, higher education is free in Norway. While healthcare is not free, there is an annual deductible above which healthcare is free. This deductible is fairly small at NOK 2,460 in 2021 ($410 in 2011 USD). Across OECD countries, Norway has the highest share of healthcare financed through government schemes and the largest per-capita spending on healthcare relating to long-term care (Cooper, 2019). As such, Norwegians in general do not seem to be susceptible to large financial costs linked to healthcare needs.

[34]Figure 1.51 shows that these patterns emerge both for short spellers and longer-term participants.

in Section 5.4, we show that those who invest for only 1–2 years perform worse than longer spellers.

3.3.3 Pensions

One may worry that the existence of pension wealth could affect individuals' desire to actively invest in the stock market out of their nonpension wealth. In principle, a rational agent should consider their overall portfolio, comprising both pension and nonpension wealth, when deciding upon their optimal portfolio allocation. If, for example, one's pension wealth is already invested in the stock market, they may invest less (or nothing at all) out of nonretirement wealth. Therefore, nonparticipation out of nonpension wealth could simply be a rational choice given existing exposure through pensions.

If pensions are to be able to explain the dynamics, the following would need to be the case: 1) the desired risky asset share out of *total* wealth changes, and individuals adjust their nonpension holdings to achieve this new goal, and/or 2) exposure to the stock market coming from pension wealth is changing at a high frequency, and individuals identify these changes and adjust their portfolio accordingly. Explaining frequent exit and (re)entry through this rebalancing channel is arguably difficult, as it requires individuals to regularly follow movements in their pension holdings and to actively rebalance accordingly. However, various papers have shown that portfolio adjustments are sluggish in both retirement and nonretirement accounts (Agnew et al., 2003; Ameriks and Zeldes, 2004; Brunnermeier and Nagel, 2008; Calvet et al., 2009a; Karlsson et al., 2009). In Section B, we provide a discussion of the Norwegian pension system and argue that the nature of the system is such that pensions are unlikely to explain the behaviors we observe.

3.3.4 Tax optimization

Could the quick exit and reentry from the stock market be due to tax optimization? Perhaps individuals choose to exit in order to reduce their tax liability in a given year. There are two possible tax margins that could be relevant here. The first is the wealth tax, whereby individuals are taxed on net wealth above a given threshold.[35] However, the majority of Norwegians do not reach the threshold. This is partly due to favorable tax treatments on certain asset classes. For example, the tax value on housing is 25% of its market value. Stocks and mutual fund holdings are given a valuation discount of

[35] In 2021, net wealth above 1.5m NOK (≈\$250,000 in 2011 USD) was taxed at 0.85% (0.7% to the municipality and 0.15% to the state). The threshold is doubled for couples.

45% (in 2021), whereas cash and deposit account holdings are not given a discount. It is therefore actually better for wealth tax purposes to retain wealth in stocks and funds rather than liquidating and holding cash. Consequently, it is very unlikely that wealth tax considerations can explain entry and exit decisions for most Norwegian households. The second relevant tax is capital gains tax. In Norway, losses made from the sale of stocks and equity funds are tax-deductible, while gains above a risk-free return are taxed. One might be worried that the quick exit we observe is because individuals are liquidating their loss-making shares to reduce their tax liabilities.[36] However, capital gains taxation in Norway is tied to the realization for each individual security, not the performance of the overall portfolio. To explain the complete exit that we observe, we would need to see every security in one's portfolio making a loss. In addition, if tax incentives are driving this behavior, we might expect to see reentrants purchasing the same stock when they return. While we do not observe specific mutual fund holdings, the Shareholder Registry provides information on direct stock ownership from 2004. We find that only 28% of directly held stocks owned just before exit are then repurchased upon reentry, meaning most reentrants are purchasing different securities. Therefore, we argue that tax-motivated selling is unlikely to drive our results.

4 Model

Our empirical results established novel patterns in the individual-level dynamics of stock market participation. This section first describes the workhorse portfolio choice model of Cocco et al. (2005) with participation costs. We then modify the model to allow agents to form beliefs over the equity premium based on realized returns, a feature motivated by a large body of literature on experience effects and memory (e.g., Kaustia and Knüpfer, 2008; Chiang et al., 2011; Malmendier and Nagel, 2011; Bordalo et al., 2020; Anagol et al., 2021). While our model embeds the participation cost story of nonparticipation, we discuss alternative theories of participation in Appendix D, namely nonstandard preferences, risks faced by households, and cultural/social factors.

[36] Using US data, Odean (1998) shows that the prevalence of selling losing stocks is highest in December, which can be linked to the end of the tax year and attempts to reduce tax liability.

4.1 Model setup

4.1.1 Preferences

Individuals are born at age t_b and die for certain by age T. They have Epstein-Zin preferences over consumption C_{it}:

$$U_{it} = \left[(1-\beta)C_{it}^{1-\frac{1}{\psi}} + \beta E_t(\pi_t U_{i,t+1}^{1-\gamma})^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

where ψ is the elasticity of intertemporal substitution, β is the subjective discount factor, and γ is the coefficient of relative risk aversion. π_t is the conditional survival probability (i.e. the probability of surviving to age $t+1$ conditional on being alive at age t). The use of Epstein-Zin preferences allows for the separation of ψ and γ, which is not possible under CRRA preferences.[37]

4.1.2 Labor market

Life is split into working age ($t \leq t_r$) and retirement ($t > t_r$), where t_r denotes retirement age. In each period, individuals receive an exogenous income Y_{it}. During working age, labor income is stochastic and depends on a deterministic function of age $f(t)$ that is calibrated to capture the hump-shaped nature of earnings during working life, as well as a transitory component u_{it} and a persistent component p_{it} modeled as a random walk.

$$\ln(Y_{it}) = f(t) + p_{it} + u_{it} \quad \text{for } t \in \{t_b, ..., t_r\},\ u_{it} \sim N(0, \sigma_u^2)$$

$$p_{it} = p_{i,t-1} + z_{it}\ ,\ z_{it} \sim N(0, \sigma_z^2)$$

We define the current level of permanent income Y_{it}^p as:

$$Y_{it}^p \equiv \exp(p_t)\exp(f(t))$$

During each year of retirement, agents receive a fraction ϕ_{ret} of their permanent income in the last year of working life. This means that upon reaching retirement age, they face no uncertainty over their future labor income.

[37] The elasticity of intertemporal substitution equals the inverse of the coefficient of relative risk aversion under CRRA utility, a restriction that does not hold in the data.

$$\ln(Y_{it}) = \ln(\phi_{ret}) + \ln(Y^p_{it_r}) = \ln(\phi_{ret}) + f(t_r) + p_{it_r} \quad \text{for} t \in \{t_r+1, ..., T\}$$

4.1.3 Financial markets and participation costs

Individuals can invest in a riskless bond with a safe net return R_f or a risky asset (stocks) with a stochastic return R_{it} given by the following:

$$1 + R_{it} = 1 + R_f + \bar{R} + \epsilon_{it} \quad \text{where } \epsilon \sim N(0, \sigma^2_\epsilon)$$

where $\bar{R}$ denotes the average equity risk premium. We do not allow agents to borrow or short sell in the model. We augment the Cocco et al. (2005) model by adding two types of stock market participation costs: an entry cost (F^0_{it}), which has to be paid at the start of any new participation spell, and a per-period cost (F^1_{it}) paid in any period where the agent chooses a positive quantity of stocks. The entry cost can reflect time and money spent figuring out how to set up an investment account, deciding on the initial portfolio, and learning fundamental investment principles. Per-period costs capture the time spent monitoring one's portfolio and deciding whether to reallocate funds, as well as any fixed account management fees (Vissing-Jørgensen, 2002). We follow Gomes and Michaelides (2005) and assume that both costs are proportional to the level of permanent income $(F^d_{it} = \bar{F}^d Y^p_{it}$ for $d \in \{0, 1\})$. We make this assumption for computational tractability. In particular, we can exploit the scale invariance of the problem and normalize the current level of permanent income Y^p_{it} to 1. However, it can be motivated by the view that participation costs reflect the opportunity cost of time.

To capture experience effects, we assume that individuals use their realized returns to update beliefs over the average equity premium $\bar{R}$. For simplicity, we assume that individuals are uncertain over whether the average premium equals its true historical value $\bar{R}_h > 0$ or a negative value $\bar{R}_l < 0$. In each period, individuals adjust their belief b_{it} that $\bar{R} = \bar{R}_h$ using Bayes rule:

$$b_{i,t+1} = \begin{cases} \frac{b_{it} f(R_{it}|\bar{R}_h)}{b_{it} f(R_{it}|\bar{R}_h) + (1-b_{it}) f(R_{it}|\bar{R}_l)} + \nu_{it} & \text{if individual } i \text{ participates at age } t \\ b_{it} + \nu_{it} & \text{otherwise} \end{cases} \tag{1.5}$$

where $\nu_{it} \sim N(0, \sigma^2_\nu)$. $f(R_{it}|\bar{R}_h)$ is the probability density function evaluated at R_{it} under

$\bar{R} = \bar{R}_h$, and similarly for $f(R_{it}|\bar{R}_l)$. Individuals are endowed with an initial belief b_{it_b} at birth.

Under this setup, when individuals experience a poor return, b_{it} is updated downwards, which means a lower expected return from participating in equities $E(R_{i,t+1})$. Our formulation aims to capture in a simple way the impact of past personal return experiences on subsequent investment behavior that has been documented in other studies. For example, Malmendier and Nagel (2011) show that a 1pp rise in experienced returns is associated with a 0.5–0.6pp rise in expected returns for the following year, which suggests that realized returns can determine participation choices through a beliefs channel. We include some noise ν_{it} in the belief formation process.[38] This means that beliefs are not completely sticky when not participating. This noise can capture various facets of behavior in a reduced-form way, such as private signals coming from external sources (e.g., from peers), limitations to retrieving memories perfectly (Azeredo da Silveira et al., 2020) or neural randomness when making choices (Fehr and Rangel, 2011).[39] We assume that individuals do not internalize the possibility of a future belief shock $\nu_{i,t+1}$ when making their optimal decision today. Shocks to beliefs are therefore completely unexpected, zero-probability events from the perspective of individuals in the model. In Section 5.3.1, we discuss how the results change under different degrees of noise including no noise at all. An implicit assumption made here is that labor income is uninformative about returns and hence it does not feature in Equation 1.5. Empirical studies have typically found the correlation between labor income and stock returns to be very close to zero (Cocco et al., 2005; Fagereng et al., 2017a), and therefore, for simplicity we assume zero correlation in the model.[40]

4.1.4 Optimization problem

Individuals choose consumption C_{it} and the risky asset share α_{it} (i.e. the share of savings allocated to the risky financial asset). The state variables are age t, cash on hand X_{it} (i.e. total resources available for consumption and saving), beliefs b_{it}, and whether one participated last year $\mathbb{1}(\alpha_{it-1} > 0)$. The latter determines whether one has to pay the

[38] To ensure that all beliefs remain within the unit interval, we truncate values to this range. When simulating the model, the standard deviation of this noise will be small, meaning minimal movement out of this interval.

[39] Noise in beliefs can relate to the wavering mechanism in Barberis et al. (2018).

[40] A different view of our setup is that it can reflect individuals having imperfect knowledge of the true returns distribution and learning about its parameters (Collin-Dufresne et al., 2016; Collard et al., 2018), or alternatively individuals learning about their own inherent ability in the stock market (Gervais and Odean, 2001; Seru et al., 2010, Linnainmaa, 2011; Anagol et al., 2021).

entry cost. Using a recursive formulation, the optimization problem can be written as follows:

$$V_t\big(X, b, \mathbb{1}(\alpha_{-1} > 0)\big) = \max_{C \geq 0, \alpha \in [0,1]} \left[(1-\beta)C^{1-\frac{1}{\psi}} + \beta E_t\Big(\pi_t V_{t+1}^{1-\gamma}\big(X', \tilde{b}', \mathbb{1}(\alpha > 0)\big)\Big)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

where

$$X' = \tilde{R}'\Big(X - C - F^0\mathbb{1}(\alpha_{-1} = 0 \ \& \ \alpha > 0) - F^1\mathbb{1}(\alpha > 0)\Big) + Y'$$

$$\tilde{R}' = 1 + R_f + \alpha(R' - R_f)$$

$$\tilde{b}' = \begin{cases} \frac{b \cdot f(R|\bar{R}_h)}{b \cdot f(R|\bar{R}_h) + (1-b) \cdot f(R|\bar{R}_l)} & \text{if } \alpha > 0 \\ b & \text{otherwise} \end{cases}$$

4.2 Calibration

Table 1.3 shows the externally calibrated parameter values used in our model simulations. Individuals are born at age $t_b = 25$ and die after age $T = 100$. For preferences, we use the median values for the coefficient of relative risk aversion ($\gamma = 5.3$) and the elasticity of intertemporal subsitution ($\psi = 0.42$) reported in Calvet et al. (2021), who estimate the cross-sectional distribution of preferences for Swedish households. We take $\beta = 0.96$ as is standard in the literature.[41] Parameter values for the income process are taken from Fagereng et al. (2017a), who estimate the process specifically for Norway. We set the high value of the average equity premium $\bar{R}_h$ equal to the historical average equity premium of 3.14% for Norway, as documented in Fagereng et al. (2017a). We assume that initial financial wealth at age t_b is drawn from a Pareto distribution, for which we take estimates of the shape and scale parameters from Fagereng et al. (2017a), who fit a Pareto distribution to the age-25 financial wealth distribution in Norway.

In the baseline simulations with beliefs, we set $\bar{F}^0 = \bar{F}^1 = 0.5\%$ ($\approx$ \$230 on average). These values are in line with the dollar per-period cost of \$250 estimated in Catherine (2021) and the portfolio adjustment cost of \$222 estimated in Choukhmane and de Silva (2022). We set the low equity premium at $\bar{R}_l = -2\%$ and allow for a small amount of noise in belief formation $\sigma_\nu = 1\%$. We provide a discussion of results under alternative parameter values in Section 5.3.

[41] This value is effectively equivalent to the median found in Calvet et al. (2021).

Table 1.3: Calibrated parameters

Parameter	Description	Value	Source
Preferences			
γ	Relative risk aversion	5.3	Calvet et al. (2021)
ψ	EIS	0.42	Calvet et al. (2021)
Institutional			
t_r	Retirement age	67	Norwegian law
π_t	Cond'l survival probabilities	-	SSB Life Tables 2010
Labor market			
$f(t)$	Deterministic wage profile	-	Fagereng et al. (2017a)
ϕ_{ret}	Replacement ratio	0.842	Fagereng et al. (2017a)
σ_z	Std. dev of permanent shock	0.110	Fagereng et al. (2017a)
σ_u	Std. dev of temporary shock	0.152	Fagereng et al. (2017a)
Financial market			
R_f	Risk-free return	0.0143	Klovland (2004)
$\bar{R}_h$	Average premium (high)	0.0314	Fagereng et al. (2017a)
σ_ϵ	Std. dev of stock return	0.238	Fagereng et al. (2017a)
μ_{x_0}	Shape of Pareto distribution	0.452	Fagereng et al. (2017a)
σ_{x_0}	Scale of Pareto distribution	5,711.7	Fagereng et al. (2017a)

Notes. This table shows the externally-calibrated parameter values used in our model simulations.

5 Results

In this section, we show the results from a model simulation of 20,000 individuals. We first consider the workhorse model with costs and quantify the size of participation costs needed to obtain reasonable short-term dynamics. Upon finding that such a model requires sizable per-period costs, we discuss how the inclusion of beliefs can produce dynamics without the need for high costs. We then test the predictions of the model using the Norwegian data.

5.1 A model without beliefs

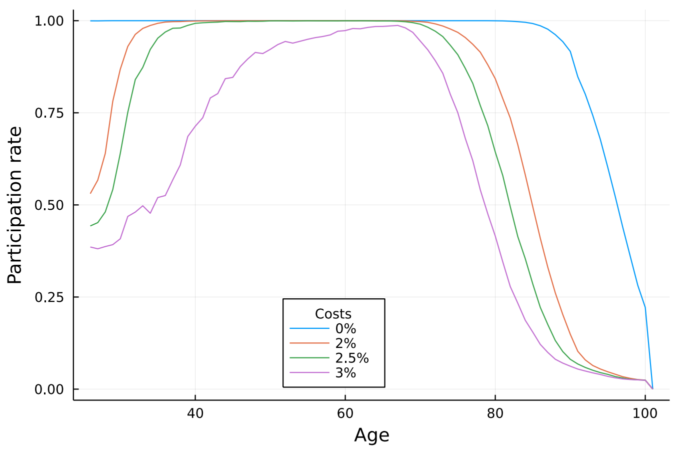
Under what conditions on participation costs can the standard model with only per-period costs and no belief heterogeneity (i.e. $b_{it} = 1\ \forall i, t$) generate short-term dynamics?[42] Figure 1.13 plots the simulated participation rates over the life cycle for different levels of per-period participation costs. In the absence of costs, individuals would invest at least a small amount in the stock market in every period in accordance with the standard Merton (1969) rule. This is because when preferences exhibit second-order risk aversion (as is the case here with Epstein-Zin preferences), risk has no first-order effect. Individuals are risk-neutral with respect to small risks, and given the positive average equity premium,

[42]In setting $b_{it} = 1\ \forall i, t$, this also means we switch off noise in beliefs ($\sigma_\nu = 0$).

zero stockholding would not be optimal (Haliassos and Bertaut, 1995). During retirement, individuals decumulate the wealth they have built up during working life to fund consumption. From around age 90, some individuals will leave the stock market because their wealth is sufficiently low to the point that they prefer to consume all of their current wealth and save nothing in either financial asset.[43]

Adding a per-period cost can delay initial entry until individuals have accumulated sufficient wealth to justify these costs. Under costs of 2% of permanent income (approximately \$900 on average), we see that the model reaches full participation quickly. The intuition for this is discussed in Gomes and Michaelides (2005) and is linked to the degree of risk aversion. A high γ generates two opposing forces: on one hand, higher risk aversion means a greater aversion to risk-taking, which directly reduces one's demand for stocks. On the other hand, it means the degree of prudence is also high, leading to strong precautionary savings motives and more wealth accumulation. As wealth increases, it will eventually be worthwhile to pay the participation costs and invest in stocks. Under $\gamma = 5.3$, this latter force dominates, which explains why the model produces full participation even with participation costs. As participation costs increase further, participation rates do not necessarily hit 100% and some nonparticipation is observed at most ages.[44]

FIGURE 1.13: Model without beliefs: simulated participation rates



Notes. This figure plots the simulated participation rate over age in the model without beliefs ($b_{it} = 1 \; \forall i, t$) for different levels of per-period participation costs. Entry costs are set to zero.

Is there much churn in participation in these models? Figure 1.14 plots the proportion of simulated spells ending within 2 years under different levels of per-period costs. To

[43]Individuals receive a certain pension income during retirement, and so not saving does not mean zero consumption in the following year.

[44]Figure 1.61 plots the average conditional risky share over the life cycle under different values of per-period costs. This share broadly falls over working age. The intuition for this follows from Jagannathan and Kocherlakota (1996), Cocco et al. (2005), and Gomes (2020). Labor income can be thought of as an implicit holding of the riskless bond given that labor income is a closer substitute to bonds than stocks (Heaton and Lucas, 1997). Over working age, income-to-financial wealth ratios decline as individuals build more wealth, leading individuals to tilt their portfolios away from stocks.

reach the prevalence of short spells observed in the data, the model requires high per-period participation costs of 2.8% of permanent income, which approximately amounts to $1,300 per annum on average. If participants are hit by an adverse shock to their income or experience a bad return, their investable wealth may be sufficiently low to no longer warrant paying the participation cost. However, this value is much higher than estimates of per-period costs in the literature (e.g., Fagereng et al., 2017a; Bonaparte et al., 2021; Catherine, 2021), and is also a large share of the median amount held in the stock market by Norwegian investors in a given year, which is $4,600. High costs are required because the savings motive in this model is so strong that small costs become redundant for the participation decision.

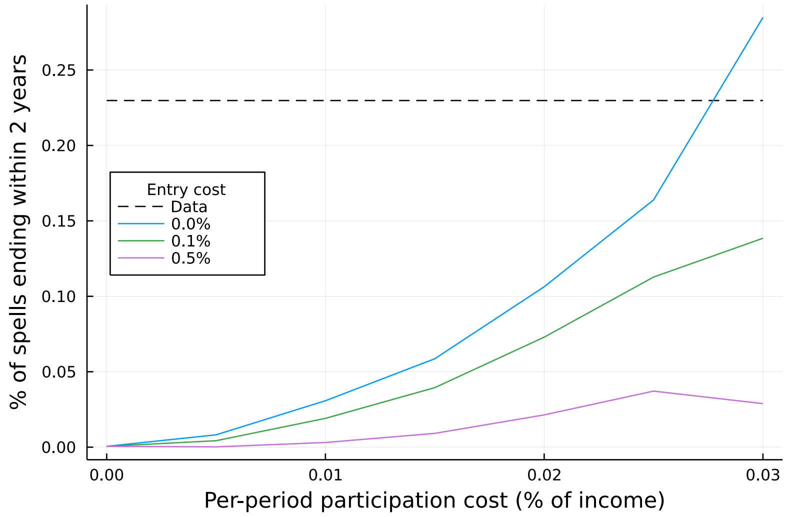
Adding even small entry costs does not help to reduce the required size of per-period costs in the model. Instead, they slow down dynamics and mean the model requires an even larger per-period cost. The intuition for this is that entry costs must be paid at the start of any new spell. Individuals recognize that exiting the market today will mean having to repay these costs again in the future should they want to reenter. As such, entry costs act as a cost of exit and lower the value of exiting today.

Figures 1.62–1.65 show the corresponding figures for the other dimensions of dynamics explored in the empirical analysis when entry costs are set to zero and per-period costs are set at 2.8%.[45] A downward-sloping hazard function for exit arises (Figure 1.62) because the longer one has been participating, the more wealth has been accumulated, which makes the costs less binding and allows them to continue participating even when faced with adverse shocks to income or returns. Reentry occurs (Figure 1.63) because exit is driven by experiencing an adverse shock and consequently having insufficient wealth to justify the participation cost. However, upon building up enough wealth, some exiters will reenter. This typically occurs very quickly owing to the fast wealth accumulation (Figure 1.64). The downward-sloping reentry hazard function (Figure 1.65) is driven by age. Given that savings motives are strong during working life, exiters during this period are certain to return. Therefore, if the agent has not returned to the stock market following a few years after exit, it must mean that they are in the retirement phase of life when individuals are drawing from their savings and thus will not reenter. Overall, the results from this analysis show that a workhorse model with income shocks and fixed participation costs can generate the short-term dynamics observed in the data provided that per-period costs are sufficiently high. However, dynamics are purely due to fluctuations in wealth in a pure cost model, and therefore, intermittent participation is concentrated in the early part of

[45] Figures 1.67–1.69 show dynamics under other values of per-period costs.

life when individuals are building up their wealth. Figure 1.66 shows that after age 40, there is virtually no exit until one reaches retirement. The pure cost model will therefore struggle to explain why middle-aged individuals may have short spells.

FIGURE 1.14: Model without beliefs: proportion of simulated spells ending within 2 years



Notes. This figure plots the proportion of simulated spells ending within 2 years in the model without beliefs ($b_{it} = 1 \; \forall i, t$) for different levels of per-period participation costs and entry costs.

5.2 A model with beliefs

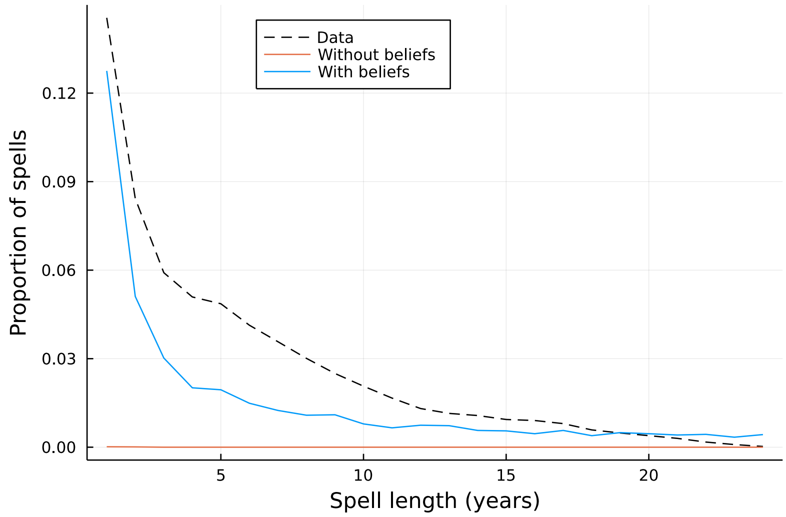
We now analyze whether the inclusion of heterogeneous beliefs and experience effects can generate short-term dynamics under smaller levels of per-period participation costs and higher levels of entry costs (both set at 0.5%). In these simulations, all individuals draw from the true returns distribution ($\bar{R} = \bar{R}_h$), but adjust their beliefs over the average equity premium following Equation 1.5.[46] Figure 1.70 plots the simulated participation rate over the life cycle. Some people are born with very pessimistic beliefs about returns. Hence, they will never enter the stock market. As a result, the participation rate is far below 100% and closer to participation rates in the data.

Figure 1.15 plots the spell length distribution for the models with and without beliefs under costs of 0.5%. Short spells are nonexistent in the pure cost model. Individuals invest at the very start of life and only exit as they approach death. Instead, the model with beliefs can produce a spell length distribution reasonably close to the data, with over 17% of spells ending within 2 years. Short spells occur because some individuals will draw poor returns, thus lowering the expected return on stocks. Some individuals will now have beliefs such that the expected return on stocks lies below the risk-free rate, and because

[46] Initial beliefs b_{it_b} are drawn from a uniform [0,1] distribution.

agents are risk averse, they will then prefer to save exclusively in bonds. The presence of participation costs generates an additional margin of exit by further reducing the net gain from stock market participation. Figure 1.71 plots the minimum wealth required to continue participating at each age for different levels of beliefs b_{it}. When the agent is certain that the average equity premium takes the higher value ($b_{it} = 1$), the wealth required is minimal. Lowering b_{it} to 0.8 raises the minimum required wealth but only marginally and mainly at the very end of life. However, if b_{it} falls further to 0.5, a belief level for which the expected equity premium is still positive, one requires a much larger level of wealth to continue participating. As a result, there is an interaction between beliefs and participation costs that can exacerbate exit. Upon experiencing a bad return, the threshold wealth an individual needs to continue investing increases, which may drive some individuals out of the market even if they believe that stocks will outperform bonds on average.

FIGURE 1.15: Model with beliefs: spell length distribution



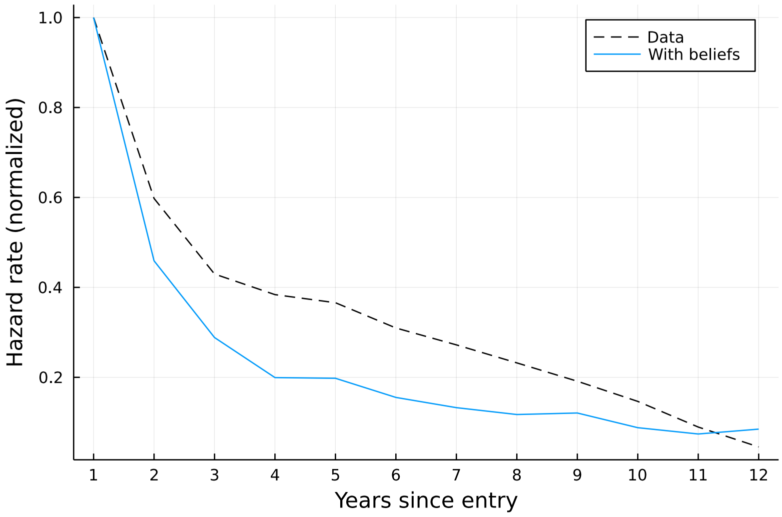
Notes. This figure plots the distribution of spell lengths in the data and the models with and without beliefs. Entry and per-period costs are both set to 0.5% of permanent income in the two models.

Figure 1.16 plots the hazard function for exit in the model with beliefs.[47] The model can generate a downward-sloping hazard function. As time spent in the stock market increases, the fact that the agent has not yet left the market must mean that they performed well in their spell thus far, and therefore, they should be optimistic about the equity premium. Consequently, one requires a very low return to undo this confidence and be driven out of the market.

Moving on to reentry, Figure 1.17 plots the distribution of the number of spells. In the model without beliefs, virtually everyone has one single, long spell lasting from birth until

[47]As short spells essentially do not occur in the model without beliefs, this model does not give a meaningful hazard function.

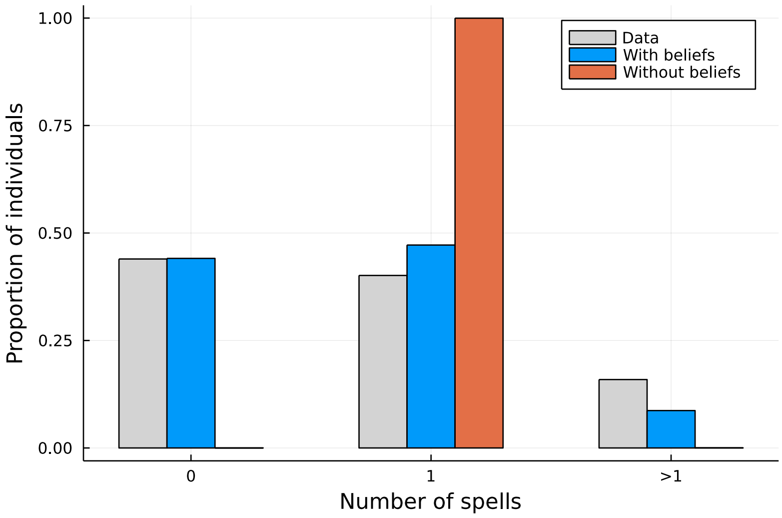
FIGURE 1.16: Model with beliefs: hazard rate for exit



Notes. This figure plots the hazard rate for exit under the model with beliefs. The hazard rate at 1 year after entry is normalized to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified (see Section 3.1.2).

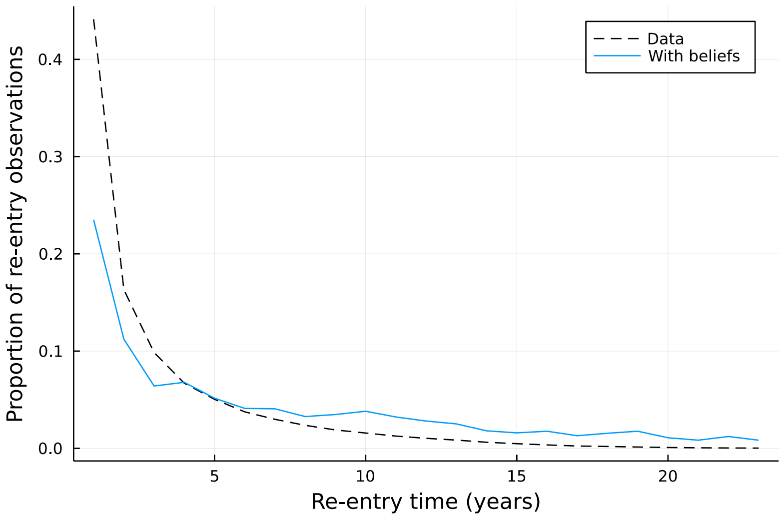
close to death, meaning that reentry does not occur and there are no "never participants". Instead, in the model with beliefs, some individuals never participate in the stock market, as they are too pessimistic about returns. This proportion of "never participants" is almost identical to the actual proportion coming from the Norwegian data. We also observe a sensible proportion of single spellers. Such individuals are a combination of people who have one long spell lasting into retirement, as well as short spellers who exit following poor returns. The model also generates reentry, albeit slightly less than is observed in the data. There are two ways in which reentry is generated in the model. First, noise in beliefs means some people may exogenously become slightly more optimistic. Note that the standard deviation of the noise shocks is set at just 1%, and thus, only nonparticipants with beliefs close to the threshold for participating (given their wealth) can be induced back into the stock market. Therefore, those who did terribly in their past spells will have weaker beliefs and will be more likely to remain out of the market indefinitely. Second, some individuals may exit because their current wealth is insufficient to warrant paying the participation costs *given their beliefs*. However, with time they may accumulate enough wealth such that it now becomes worthwhile to reenter. Figure 1.18 plots the simulated reentry times. While the model does not generate as many 1-year reentry observations as in the data, 1-year reentry remains the modal outcome. The model produces quick reentry because the noise in beliefs can drive recent exiters with beliefs close to the participation threshold back into the market. We obtain more longer-term reentry relative to the data because of the second channel for reentry. Some exiters have beliefs such that the expected equity premium is positive, but with participation costs, they need to accumulate more wealth to justify participating, and this can take some time. We obtain a downward-sloping hazard function reasonably close to the data. The intuition for this result is that those who have not reentered after many years will typically have pessimistic beliefs. Such individuals will likely not be drawn back in through the optimistic swings in beliefs coming through the noise term. Hence, their reentry is less likely.

FIGURE 1.17: Model with beliefs: number of spells



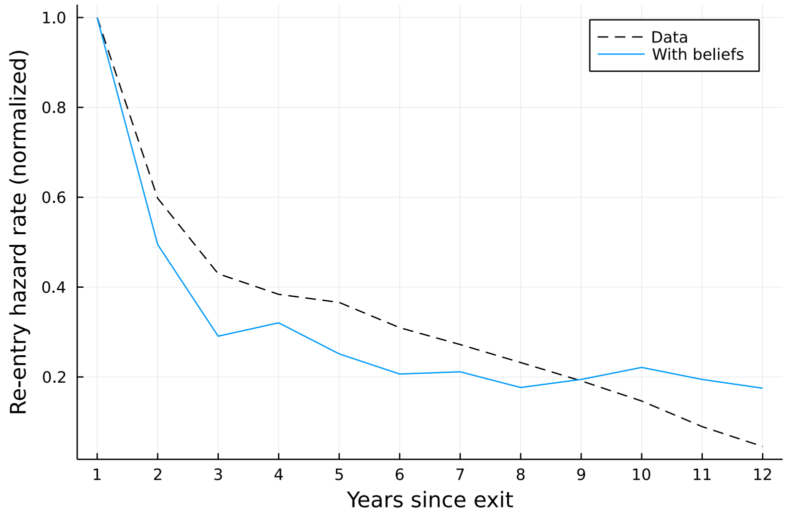
Notes. This figure plots the distribution of the number of spells in the simulated population in the models with and without beliefs. The empirical distribution for the Norwegian population is also shown (see Figure 1.9).

FIGURE 1.18: Model with beliefs: reentry times



Notes. This figure plots the distribution of reentry times in the model with beliefs. The empirical proportion from the Norwegian data is also shown (see Figure 1.11).

FIGURE 1.19: Model with beliefs: hazard rate for reentry



Notes. This figure plots the hazard rate for reentry under the model with beliefs. The hazard rate at 1 year after exit is normalized to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified (see Sections 3.1.2 and 3.2.3).

5.3 Sensitivity to different parameter values

5.3.1 Degree of noise in beliefs

We consider different values for the standard deviation of belief shocks $\sigma_\nu \in \{0\%, 1\%, 2\%\}$ in Figures 1.72–1.78. A lower degree of noise in belief formation reduces the participation rate at all ages (Figure 1.72) because the presence of noise can help drive some individuals back into the stock market. In the extreme case of zero noise, beliefs are completely sticky for nonparticipants. The case of zero noise does not mean no reentry. Some individuals exit because they do not currently have sufficient wealth to justify the participation costs given their beliefs. Upon accumulating more wealth, they may eventually reenter. However, the degree of reentry will be less compared to the case with noise (Figure 1.76). In addition, noise can generate quicker reentry. Without noise, the only way individuals can reenter is by accumulating more wealth such that they can now justify the costs, and this can sometimes take time. With noise in beliefs, people may become a bit more optimistic, which can induce quicker reentry because they need a smaller amount of wealth to warrant paying the per-period costs and reenter (as depicted in Figure 1.71). The presence of noise can also lead to more short spells (Figure 1.74) because individuals can experience negative swings in optimism, which can drive those with beliefs close to the threshold out of the market. Therefore, the presence of some noise in belief formation can help to better match the degree of short spells and reentry observed in the data.

5.3.2 Participation costs

In the baseline simulations, we set per-period and entry costs to 0.5% of permanent income. However, the model does not require the existence of participation costs to generate short-term dynamics (Figures 1.79–1.85). Reducing costs right down to zero means insufficient wealth is no longer a driver of exit. Instead, beliefs become the sole reason to exit.[48] Figure 1.81 shows that the prevalence of short spells is actually higher under zero costs than under 0.5% costs. Furthermore, Figure 1.83 shows that a higher share of individuals have multiple spells in the model without costs. This is because entry costs deter exit, and therefore, its removal encourages more temporary exit. Overall, the model does not rely on participation costs to generate reasonable degrees of short-term dynamics.

5.4 Supporting evidence for the model mechanisms

Our model gives predictions regarding who should exit from or reenter the stock market. First, in the model with beliefs, those who stop participating soon after entry should have weaker average performance relative to those who stayed in the market for longer. At the aggregate level, as discussed in Section 3.1.1, the prevalence of short spells in mutual funds rises during stock market downturns (Figure 1.31a). We also see that average risky shares fall during downturns (Figure 1.86).[49] At the individual level, Figure 1.20 suggests that short spellers do poorly relative to longer spellers. We measure performance by computing the proportion of exiters of different spell lengths reporting only taxable gains from the sale of stocks and equity funds (Figure 1.20a) or only losses (Figure 1.20b) in their exit year.[50] The unconditional probability of reporting only gains is 30% for short spellers compared to around 40% for those participating for longer. Similarly, the unconditional probability of reporting only losses for short spellers (≈28%) is twice that of longer spellers

[48] This is the case other than the years just before death. During this period late in life, some people may not save at all (in either asset), as they have very low values of wealth. They prefer to consume everything because they know they will receive a pension income in the next period. In other periods of life, precautionary savings motives are sufficiently high that everyone would have enough wealth to warrant saving.

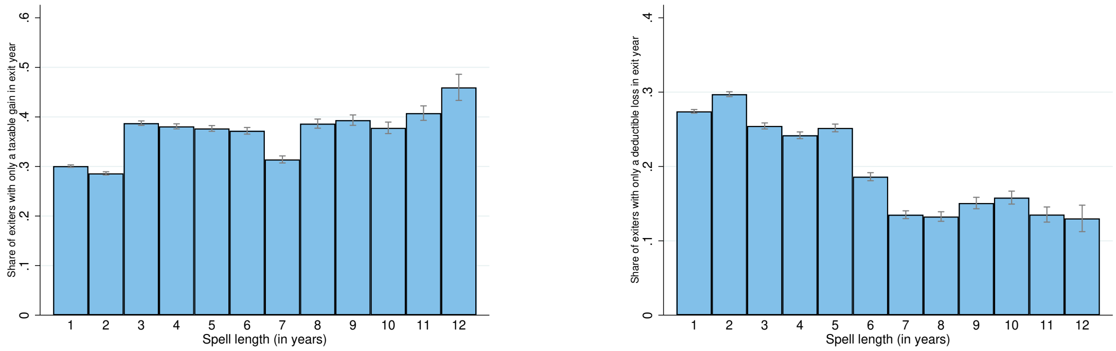
[49] It is important to note that part of this is likely reflecting passive drops in portfolio values rather than active changes in portfolio holdings. As we only see the total market value held in funds at the end of the year and thus do not observe the specific funds that individuals own, we cannot confidently separate these two effects.

[50] For this analysis, we restrict attention to exiters who entered from 2006 onward because of changes in the Norwegian tax system that make it difficult to interpret the tax record variables prior to this point. Since 2006, individuals are only taxed on capital gains above a risk-free return. However, before 2006 the taxable amount depended on the share's proportion of retained taxed capital, and thus, it may necessarily not be linked to achieving a high/low return relative to a risk-free asset. Taxed capital refers to undistributed income that has been previously subject to tax at the company level. Focusing only on exiters who entered from 2006 onward aids with the interpretation of the tax variables because these individuals would be subject to the "new" tax system based on returns relative to a risk-free rate.

(≈13–15%).[51] While one may be concerned that the higher prevalence of losses among short spellers reflects a liquidity need that forces them to liquidate at a loss, the discussion in Section 3.3.1 suggests that these shocks cannot explain the quick exit observed in the data.[52]

FIGURE 1.20: Performance of exiters by spell length

(A) Report only gains (B) Report only losses



Notes. This figure shows the performance of exiters by spell length based on records of taxable gains and tax-deductible losses in the income tax data. In panel (A), we plot the proportion of exiters of a given spell length reporting only gains from the sale of stocks and funds (computed as the sum of items TR 3.1.8, TR 3.1.9, and TR 3.1.10 in the tax records) in their exit year. In panel (B), we plot the proportion of exiters reporting only losses (computed as the sum of items TR 3.3.8, TR 3.3.9, and TR 3.3.10). We use exiters who enter from 2006 onward in these plots.

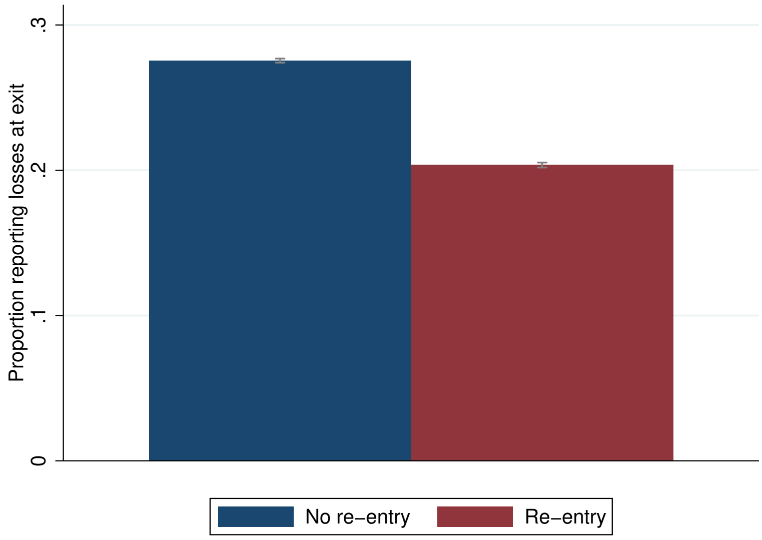
Second, in terms of reentry, the model predicts those who end up returning to the stock market should on average have done better in their prior spell compared to those who chose to remain out of the market. Figure 1.21 shows this to be the case. About 27% of those who do not reenter experience losses compared to 20% for individuals who reenter.[53]

[51] Figure 1.48 plots the corresponding figures based on reporting any gains or any losses rather than only gains or losses. We obtain broadly similar findings.

[52] We obtain similar conclusions when looking at the size of gains and losses relative to risky financial wealth (Figure 1.87).

[53] Figure 1.88 also shows that exiters in Norway who report only taxable gains are about 11% (3pps) more likely to reenter than those who report only losses.

Figure 1.21: Prevalence of losses by reentry status



Notes. This figure plots the share of people experiencing losses separately for those who reentered into the stock market and those who did not. We use exiters who enter from 2006 onward in these plots.

6 Conclusion

While there is a large body of literature that studies why many individuals do not invest in stocks, much less is known about the dynamics of stock market participation by retail investors. How long do individuals stay in the stock market for? Is the probability of exit a function of time since entry? Do individuals reenter after exit, and if so, when? Using Norwegian administrative data, we document new facts regarding the exit and reentry decisions of individual investors. The unifying message from these facts is that short, multiple spells in the stock market are common. We show that the workhorse portfolio choice model needs high per-period participation costs to generate these patterns. Extending the model to allow for experience effects à la Malmendier and Nagel (2011), whereby individuals adjust their expected stock returns based on realized returns, is able to produce short-term dynamics without the need for high participation costs.

Our findings leave various avenues for future research. First, the current setup of the model does not contain aggregate signals from which nonparticipants can learn about stock returns. Extending the model to allow for this can help to explain patterns in the data such as why entry rates tend to drop during stock market downturns. Second, what does noise in belief formation in the model exactly represent? One possible explanation is imperfect memory recollection. While the neuroscience and psychology literatures have established that memory is imperfect, there is little empirical evidence directly testing imperfect memory in the context of financial markets. Further work trying to see how well former participants recall their past return experiences and what biases they are prone to would thus help to establish whether noisy memory is a feature of investor behavior. Third, the Norwegian data allows one to identify family networks. It would be informative to directly test whether peer effects play a role in determining spell length and reentry. Last, our finding that a large share of the population have short-lived spells in the stock market can have important implications for wealth accumulation. If individuals are liquidating their entire stockholdings soon after entry, they are not staying in the stock market for long enough to earn the average equity premium, which can hurt their wealth accumulation going forward. This is particularly the case when individuals are permanently scarred by past adverse returns. Policies should therefore not solely focus on encouraging entry into the stock market. They also need to address the fact that many people quickly exit. Identifying policies that can achieve longer-term participation is important, particularly as individuals of lower wealth and education appear to be more prone to such intermittent spells in the stock market.

References

Aastveit, K. A., T. M. Fastbø, E. Granziera, K. S. Paulsen, and K. N. Torstensen (2020): "Nowcasting Norwegian household consumption with debit card transaction data," Discussion paper, Norges Bank Working Paper.

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2020): "Overreaction in expectations: Evidence and theory," *Available at SSRN*.

Agnew, J., P. Balduzzi, and A. Sundén (2003): "Portfolio Choice and Trading in a Large 401(k) Plan," *American Economic Review*, 93(1), 193–215.

Alvarez, F. E., K. Borovičková, and R. Shimer (2021): "Consistent Evidence on Duration Dependence of Price Changes," Working Paper 29112, National Bureau of Economic Research.

Ameriks, J., and S. P. Zeldes (2004): "How do household portfolio shares vary with age," Discussion paper, Columbia University.

Anagol, S., V. Balasubramaniam, and T. Ramadorai (2021): "Learning from noise: Evidence from India's IPO lotteries," *Journal of Financial Economics*, 140(3), 965–986.

Ang, A., G. Bekaert, and J. Liu (2005): "Why stocks may disappoint," *Journal of Financial Economics*, 76(3), 471–508.

Azeredo da Silveira, R., Y. Sung, and M. Woodford (2020): "Optimally imprecise memory and biased forecasts," Discussion paper, National Bureau of Economic Research.

Bach, L., L. E. Calvet, and P. Sodini (2020): "Rich Pickings? Risk, Return, and Skill in Household Wealth," *American Economic Review*, 110(9), 2703–47.

Barber, B. M., and T. Odean (2000): "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *The Journal of Finance*, 55(2), 773–806.

——— (2001): "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment," *The Quarterly Journal of Economics*, 116(1), 261–292.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2018): "Extrapolation and bubbles," *Journal of Financial Economics*, 129(2), 203–227.

Basten, C., A. Fagereng, and K. Telle (2016): "Saving and Portfolio Allocation Before and After Job Loss," *Journal of Money, Credit and Banking*, 48(2-3), 293–324.

Behrman, J. R., O. S. Mitchell, C. K. Soo, and D. Bravo (2012): "How Financial Literacy Affects Household Wealth Accumulation," *American Economic Review*, 102(3), 300–304.

Benhabib, J., A. Bisin, and M. Luo (2019): "Wealth Distribution and Social Mobility in the US: A Quantitative Approach," *American Economic Review*, 109(5), 1623–47.

Benhabib, J., A. Bisin, and S. Zhu (2011): "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 79(1), 123–157.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2007): "Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated," *The Journal of Finance*, 62(5), 2123–2167.

Bonaparte, Y., G. M. Korniotis, and A. Kumar (2021): "Income Risk and Stock Market Entry/Exit Decisions," Discussion paper, CEPR Discussion Paper No. DP15370.

Bordalo, P., N. Gennaioli, and A. Shleifer (2020): "Memory, Attention, and Choice," *The Quarterly Journal of Economics*, 135(3), 1399–1442.

Brandsaas, E. E. (2021): "Household Stock Market Participation and Exit: The Role of Homeownership," .

Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner (2008): "Neighbors Matter: Causal Community Effects and Stock Market Participation," *The Journal of Finance*, 63(3), 1509–1531.

Brunnermeier, M. K., and S. Nagel (2008): "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-evidence on Individuals," *American Economic Review*, 98(3), 713–36.

Calvet, L. E., J. Y. Campbell, F. J. Gomes, and P. Sodini (2021): "The Cross-Section of Household Preferences," Discussion paper, Working Paper.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2007): "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115(5), 707–747.

——— (2009a): "Fight or Flight? Portfolio Rebalancing by Individual Investors," *The Quarterly Journal of Economics*, 124(1), 301–348.

——— (2009b): "Measuring the Financial Sophistication of Households," *American Economic Review*, 99(2), 393–98.

Campbell, J. Y. (2006): "Household Finance," *The Journal of Finance*, 61(4), 1553–1604.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–251.

Cao, H. H., T. Wang, and H. H. Zhang (2005): "Model Uncertainty, Limited Market Participation, and Asset Prices," *The Review of Financial Studies*, 18(4), 1219–1251.

Carroll, C. D., B.-K. Rhee, and C. Rhee (1994): "Are There Cultural Effects on Saving? Some Cross-Sectional Evidence," *The Quarterly Journal of Economics*, 109(3), 685–699.

——— (1999): "Does Cultural Origin Affect Saving Behavior? Evidence from Immigrants," *Economic Development and Cultural Change*, 48(1), 33–50.

Catherine, S. (2021): "Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle," *The Review of Financial Studies*, 35(9), 4016–4054.

Chiang, Y.-M., D. Hirshleifer, Y. Qian, and A. E. Sherman (2011): "Do Investors Learn from Experience? Evidence from Frequent IPO Investors," *The Review of Financial Studies*, 24(5), 1560–1589.

Choi, J. J., and A. Z. Robertson (2020): "What Matters to Individual Investors? Evidence from the Horse's Mouth," *The Journal of Finance*, 75(4), 1965–2020.

Choukhmane, T., and T. de Silva (2022): "What Drives Investors' Portfolio Choices? Separating Risk Preferences from Frictions," Discussion paper, Working paper.

Christiansen, C., J. S. Joensen, and J. Rangvid (2015): "Understanding the effects of marriage and divorce on financial investments: the role of background risk sharing," *Economic Inquiry*, 53(1), 431–447.

Chuang, Y., and L. Schechter (2015): “Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results,” *Journal of Development Economics*, 117, 151–170.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout (2005): “Consumption and Portfolio Choice over the Life Cycle,” *The Review of Financial Studies*, 18(2), 491–533.

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon (2018): “Ambiguity and the historical equity premium,” *Quantitative Economics*, 9(2), 945–993.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016): “Parameter Learning in General Equilibrium: The Asset Pricing Implications,” *American Economic Review*, 106(3), 664–98.

Constantinides, G. (1990): “Habit Formation: A Resolution of the Equity Premium Puzzle,” *Journal of Political Economy*, 98(3), 519–43.

Cooper, J. (2019): “How does UK healthcare spending compare with other countries,” *Office for National Statistics*.

Dohmen, T., H. Lehmann, and N. Pignatti (2016): “Time-varying individual risk attitudes over the Great Recession: A comparison of Germany and Ukraine,” *Journal of Comparative Economics*, 44(1), 182–200.

Epstein, L. G., and T. Wang (1994): “Intertemporal Asset Pricing under Knightian Uncertainty,” *Econometrica*, 62(2), 283–322.

Epstein, L. G., and S. E. Zin (1990): “‘First-order’ risk aversion and the equity premium puzzle,” *Journal of Monetary Economics*, 26(3), 387–407.

Fagereng, A., C. Gottlieb, and L. Guiso (2017a): “Asset Market Participation and Portfolio Choice over the Life-Cycle,” *The Journal of Finance*, 72(2), 705–750.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020): “Heterogeneity and Persistence in Returns to Wealth,” *Econometrica*, 88(1), 115–170.

Fagereng, A., L. Guiso, and L. Pistaferri (2017b): “Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk,” *The Review of Economic Studies*, 85(1), 437–474.

Fagereng, A., M. B. Holm, B. Moll, and G. Natvik (2019): “Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains,” Working Paper 26588, National Bureau of Economic Research.

Fehr, E., and A. Rangel (2011): "Neuroeconomic Foundations of Economic Choice–Recent Advances," *Journal of Economic Perspectives*, 25(4), 3–30.

Frazzini, A., and O. A. Lamont (2008): "Dumb money: Mutual fund flows and the cross-section of stock returns," *Journal of Financial Economics*, 88(2), 299–322.

Fredriksen, D., and E. Halvorsen (2019): "Beregninger av pensjonsformue," Discussion paper, Statistics Norway.

Fredriksen, D., and N. M. Stølen (2018): "Reform av offentlig tjenestepensjon," Discussion paper, Statistics Norway.

Fuchs-Schündeln, N., P. Masella, and H. Paule, Paludkiewicz (2020): "Cultural Determinants of Household Saving Behavior," *Journal of Money, Credit and Banking*, 52(5), 1035–1070.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016): "The Dynamics of Inequality," *Econometrica*, 84(6), 2071–2111.

Gervais, S., and T. Odean (2001): "Learning to Be Overconfident," *The Review of Financial Studies*, 14(1), 1–27.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

Gomes, F. (2020): "Portfolio Choice Over the Life Cycle: A Survey," *Annual Review of Financial Economics*, 12(1), 277–304.

Gomes, F., M. Haliassos, and T. Ramadorai (2021): "Household Finance," *Journal of Economic Literature*, 59(3), 919–1000.

Gomes, F., and A. Michaelides (2005): "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence," *The Journal of Finance*, 60(2), 869–904.

Greenwood, R., and S. Nagel (2009): "Inexperienced investors and bubbles," *Journal of Financial Economics*, 93(2), 239–258.

Grinblatt, M., and M. Keloharju (2009): "Sensation Seeking, Overconfidence, and Trading Activity," *The Journal of Finance*, 64(2), 549–578.

Gruber, M. J. (1996): "Another Puzzle: The Growth in Actively Managed Mutual Funds," *The Journal of Finance*, 51(3), 783–810.

Guin, B. (2017): "Culture and household saving," Working Paper Series 2069, European Central Bank.

Guiso, L., M. Haliassos, T. Jappelli, and S. Claessens (2003a): "Household Stockholding in Europe: Where Do We Stand and Where Do We Go?," *Economic Policy*, 18(36), 125–170.

Guiso, L., P. Sapienza, and L. Zingales (2003b): "People's opium? Religion and economic attitudes," *Journal of Monetary Economics*, 50(1), 225–282.

——— (2006): "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 20(2), 23–48.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667–686.

Haliassos, M., and C. C. Bertaut (1995): "Why do so Few Hold Stocks?," *The Economic Journal*, 105(432), 1110–1129.

Haliassos, M., T. Jansson, and Y. Karabulut (2017): "Incompatible European Partners? Cultural Predispositions and Household Financial Behavior," *Management Science*, 63(11), 3780–3808.

Haliassos, M., and A. Michaelides (2003): "Portfolio Choice and Liquidity Constraints," *International Economic Review*, 44(1), 143–177.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

Heaton, J., and D. Lucas (1997): "Market Frictions, Savings Behavior, and Portfolio Choice," *Macroeconomic Dynamics*, 1(1), 76–101.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271–320.

Hong, H., J. D. Kubik, and J. C. Stein (2004): "Social Interaction and Stock-Market Participation," *The Journal of Finance*, 59(1), 137–163.

Hong, H., and J. C. Stein (2007): "Disagreement and the Stock Market," *Journal of Economic Perspectives*, 21(2), 109–128.

Honoré, B. E. (1993): "Identification Results for Duration Models with Multiple Spells," *Review of Economic Studies*, 60(1), 241–46.

Hubmer, J., P. Krusell, and A. A. Smith. (2021): "Sources of US Wealth Inequality: Past, Present, and Future," *NBER Macroeconomics Annual*, 35, 391–455.

Hurst, E., M. C. Luoh, F. P. Stafford, and W. G. Gale (1998): "The Wealth Dynamics of American Families, 1984-94," *Brookings Papers on Economic Activity*, 1998(1), 267–337.

Ippolito, R. A. (1992): "Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry," *The Journal of Law and Economics*, 35(1), 45–70.

Jagannathan, R., and N. Kocherlakota (1996): "Why Should Older People Invest Less in Stocks Than Younger People?," *Quarterly Review*.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–291.

Karlsson, N., G. Loewenstein, and D. Seppi (2009): "The ostrich effect: Selective attention to information," *Journal of Risk and uncertainty*, 38(2), 95–115.

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

Kaustia, M., and S. Knüpfer (2012): "Peer performance and stock market entry," *Journal of Financial Economics*, 104(2), 321–338.

Kiefer, N. M. (1988): "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26(2), 646–679.

Klovland, J. T. (2004): "Bond markets and bond yields in Norway 1820–2003," in *Historical Monetary Statistics for Norway 1819-2003*, ed. by Øyvind Eitrheim, J. T. Klovland, and J. F. Qvigstad, no. 35 in Norges Bank Occasional Papers. Norges Bank.

Lancaster, T. (1979): "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47(4), 939–956.

Linnainmaa, J. T. (2011): "Why Do (Some) Households Trade So Much?," *The Review of Financial Studies*, 24(5), 1630–1666.

Lusardi, A., and O. S. Mitchell (2011): "Financial Literacy and Planning: Implications for Retirement Well-being," in *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*. Oxford University Press, Oxford.

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

——— (2015): " Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 131(1), 53–87.

MALMENDIER, U., AND J. A. WACHTER (2021): "Memory of Past Experiences and Economic Decisions," *Available at SSRN 4013583*.

MANKIW, N., AND S. P. ZELDES (1991): "The consumption of stockholders and non-stockholders," *Journal of Financial Economics*, 29(1), 97–112.

MERTON, R. C. (1969): "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *The Review of Economics and Statistics*, 51(3), 247–257.

MERTON, R. C. (1971): "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3(4), 373–413.

MEYER, B. D. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757–782.

NAKAMURA, E., AND J. STEINSSON (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.

ODEAN, T. (1998): "Are Investors Reluctant to Realize Their Losses?," *The Journal of Finance*, 53(5), 1775–1798.

OECD (2009): "OECD Private Pensions Outlook 2008," Discussion paper, OECD.

OFEK, E., AND M. RICHARDSON (2003): "DotCom Mania: The Rise and Fall of Internet Stock Prices," *The Journal of Finance*, 58(3), 1113–1137.

OSILI, U. O., AND A. PAULSON (2008): "What Can We Learn about Financial Access from U.S. Immigrants? The Role of Country of Origin Institutions and Immigrant Beliefs," *The World Bank Economic Review*, 22(3), 431–455.

PAGEL, M. (2018): "A news-utility theory for inattention and delegation in portfolio choice," *Econometrica*, 86(2), 491–522.

POTERBA, J. M., AND A. A. SAMWICK (1997): "Household Portfolio Allocation Over the Life Cycle," Working Paper 6185, National Bureau of Economic Research.

QUIGGIN, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3(4), 323–343.

ROUTLEDGE, B. R., AND S. E. ZIN (2010): "Generalized Disappointment Aversion and Asset Prices," *The Journal of Finance*, 65(4), 1303–1332.

Samuelson, P. A. (1969): "Lifetime Portfolio Selection By Dynamic Stochastic Programming," *The Review of Economics and Statistics*, 51(3), 239–246.

Schildberg-Hörisch, H. (2018): "Are Risk Preferences Stable?," *Journal of Economic Perspectives*, 32(2), 135–54.

Segal, U., and A. Spivak (1990): "First order versus second order risk aversion," *Journal of Economic Theory*, 51(1), 111–125.

Seru, A., T. Shumway, and N. Stoffman (2010): "Learning by Trading," *The Review of Financial Studies*, 23(2), 705–739.

Shefrin, H., and M. Statman (1985): "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence," *The Journal of Finance*, 40(3), 777–790.

Shiller, R. J., S. Fischer, and B. M. Friedman (1984): "Stock Prices and Social Dynamics," *Brookings Papers on Economic Activity*, 1984(2), 457–510.

Vissing-Jørgensen, A. (2002): "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures," Working Paper 8884, National Bureau of Economic Research.

——— (2003): "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," *NBER Macroeconomics Annual*, 18, 139–194.

Xavier, I. (2021): "Wealth inequality in the US: the role of heterogeneous returns," *Available at SSRN 3915439*.

Appendix

A Variable construction

Here, we describe the steps undertaken to translate the tax records into consistent measures of wealth by broad asset class. TR x.y denotes item x.y in the tax records based on 2018 item codings by the Norwegian Tax Administration (Skatteetaten). Note that while tax values are reported in the raw data, we translate these values into market values for our analysis. For financial wealth, we create the following subclasses:

- Cash and deposits are computed as the sum of deposits in Norwegian banks (TR 4.1.1), cash (TR 4.1.3), deposits in foreign banks (TR 4.1.9), and (from 2017 onward) cash holdings in share savings accounts (TR 4.1.8.6).
- Directly held listed stocks are given by the value of listed Norwegian shares and equity certificates, bonds, etc. in the Norwegian Central Securities Depository (TR 4.1.7).
- Directly held unlisted stocks are given by capital in unlisted shares, share savings accounts, and securities not listed in the Norwegian Central Securities Depository (TR 4.1.8).
- Stock mutual fund holdings are given by the value of the share component in holdings of securities funds (TR 4.1.4) plus (from 2017 onward) equity holdings in share savings accounts (TR 4.1.8.5).
- Money market/bond funds are given by the value of the interest component in holdings of securities funds (TR 4.1.5).
- Financial wealth held abroad is given by other taxable capital abroad such as foreign shares, outstanding claims, bonds, and endowment insurance (TR 4.6.2).
- Other financial assets are the sum of outstanding receivables in Norway (TR 4.1.6), the share of capital in housing cooperatives or jointly-owned property (TR 4.5.3), pension insurance and life insurance (TR 4.5.1 + TR 4.5.2), and other taxable capital, such as cryptocurrency (TR 4.5.4).

Real wealth can be decomposed into the follow:

- Housing wealth is the sum of housing owned through housing cooperatives (TR 4.3.2.2) and self-owned property (TR 4.3.2.1 + TR 4.3.2.3).

- Other real wealth is the sum of boats (TR 4.2.4), cars (TR 4.2.5), caravans (TR 4.2.6), holiday homes (TR 4.3.3.1 + TR 4.3.2.3), other real estate (TR 4.3.4 + TR 4.3.5 + TR 4.3.2.3), home contents and movable property (TR 4.2.3), fixtures and other business assets (TR 4.4.1 + TR 4.4.2 + TR 4.4.3 + TR 4.4.4), and real wealth abroad (TR 4.6.1 + TR 4.3.6.1).

B The Norwegian pension system

There are three main components to the Norwegian pension system. First is the National Insurance Scheme ("*folketrygden*"), which is the basic public pension scheme. It ensures that everyone receives a minimum pension income. Furthermore, workers are guaranteed a supplement that is proportional to their income during working age.[54] The system is defined-benefit in nature, so citizens face no stock market exposure through it. As such, the decisions to exit and enter the stock market cannot be attributed to portfolio rebalancing between private accounts and public pension wealth.

Second, there are occupational pensions. Public occupational pensions are also defined-benefit schemes. Hence, there is no stock market exposure through them.[55] Private sector occupational pensions operate differently. Until 2001, only defined-benefit pensions existed. While defined-contribution pensions, for which the pension benefit depends on how well the contributions are invested, have been allowed since 2001, they did not gain momentum until 2006 when occupational pensions were made mandatory by law. Indeed, before 2006 occupational pensions were mainly provided by larger employers (OECD, 2009).[56] One may be concerned that because private sector defined-contribution occupa-

[54]Under the current system, 18.1% of wages in each year of employment up to a certain ceiling is transferred to a pension account. This pension income is then indexed to nominal wage growth. Upon retirement, the accumulated amount is not given as a lump sum. Instead, an annual sum is given based on the expected number of years to be spent as a pensioner, which itself depends on when the individual starts withdrawing from their pension and life expectancy. While there are some differences based on year of birth, the overall premise of pensionable income being linked to employment earnings still holds. For further details, see Fagereng et al. (2019) and Fredriksen and Halvorsen (2019).

[55]Until 2020, the public occupational pension scheme was such that workers were entitled to the maximum pension after 30 years of service and can receive a pension equal to 66% of their pension base (final salary converted into a full-time equivalent) before adjustments for life expectancy. However, from 2020 occupational pension earnings became similar to that in the National Insurance Scheme, in particular having a share of earnings each year be accumulated in a pension pot. However, this remained a defined-benefit system. For further details on public occupational pensions and the reforms, see Fredriksen and Stølen (2018).

[56]As of 2018, 90% of private sector employees are under a defined-contribution pension (Fredriksen and Halvorsen, 2019).

tional pensions have some exposure to the stock market, this could influence choices made in nonretirement investment accounts. However, Figure 1.7 shows that short spells in the stock market are not exclusive to the post-2006 period.

Third, individuals may have personal private pensions that they invest in. As payments into an Individual Pension Scheme (IPS) in Norway are tax deductible up to a certain limit, one can infer from the tax records whether an individual holds such pensions.[57] Figure 1.58 provides a time series of participation in private pension accounts separately for the whole population and the subset of the population aged 60 and under (who are unlikely to have drawn from such pensions yet). In either case, the participation rates are in single digits, indicating that the vast majority of the population do not hold such accounts. To further ease concerns, we plot the proportion of exiters of different spell lengths who hold private pensions as of their exit year. If these schemes were driving our short spell result, we might expect to see a greater prevalence of private pensions among short spellers. However, Figure 1.59 shows the opposite. We also reproduce our spell length histogram but exclude any individual who at any point in the sample holds a private pension account. Figure 1.60 shows that our results are robust to this. We therefore believe that pension holdings cannot explain the short-term dynamics we observe.

C Further details on the Alvarez et al. (2021) GMM estimator

The Alvarez et al. (2021) GMM estimator is based on the following environment. There is a proportional hazards data generating process for durations $d \in \{\underline{D}, ..., \bar{D}\}$, where $h_i(d) = \theta_i b_d$. θ_i is the time-invariant frailty parameter specific to individual i and captures individual heterogeneity in hazard rates. b_d is the baseline hazard at duration d and is assumed to be common across individuals. The objective is to obtain an estimate of b_d, as this reflects true duration dependence rather than unobserved heterogeneity. Individual i experiences K^i spells, for which the measured duration of spells is $\zeta^i = \{\zeta_0^i, \zeta_1^i, ..., \zeta_{K^i}^i\}$. Note that measured duration is not necessarily equal to the true length of the spell because of censoring. Assume that the spells $\zeta = (\zeta_0, \zeta_1, ..., \zeta_K)$ are drawn from a proportional

[57] There are two relevant variables in the tax data. TR 3.3.5 records the deductible amount from payments into an IPS, while TR 4.5.1 indicates capital in an Individual Pension Account (IPA). Note that IPAs were replaced by the IPS in 2006, from which point new money could not be placed into one's existing IPA, and new IPAs could not be opened. We consider an individual to be a private pension contributor if they report a positive value for either of these two variables, either in the current year or in any past year.

hazards model with a baseline hazard $\mathbf{b}_0$. Defining

$$f^{[b]}_{d_1,d_2}(\zeta;\mathbf{b}) \equiv \sum_{(j,k):1\leq j\leq k\leq K} \left(b_{d_2}\mathbb{1}_{\zeta_j=d_1,\zeta_k\geq d_2} - b_{d_1}\mathbb{1}_{\zeta_j=d_2,\zeta_k\geq d_1}\right)$$

then $\mathbb{E}[f^{[b]}_{t_1,t_2}] = 0 \;\forall\; \underline{D} \leq d_1 < d_2 \leq \bar{D}$ if and only if $\mathbf{b} = \lambda\mathbf{b}_0$ for some $\lambda > 0$. This gives $\frac{\tilde{D}(\tilde{D}+1)}{2}$ moment conditions, where $\tilde{D} \equiv \bar{D} - \underline{D}$. It is important to note that under this procedure, we recover the baseline hazards $\mathbf{b}$ up to a multiplicative constant, and so we normalize $b_1 = 1$. To estimate $\mathbf{b}_0$:

$$\hat{\mathbf{b}}_0 = \arg\min_{\mathbf{b}} \Big(\frac{1}{N}\sum_{i=1}^{N} f^{[b]}_{d_1,d_2}(\zeta^i;\mathbf{b})\Big)^T W \Big(\frac{1}{N}\sum_{i=1}^{N} f^{[b]}_{d_1,d_2}(\zeta^i;\mathbf{b})\Big)$$

where W is a positive definite weighting matrix. We use two-step feasible GMM à la Hansen (1982). In the first step, we use the identity matrix as the weighting matrix. In the second step, we take our estimates from the first step, $\mathbf{b}_0^{(1)}$, and use $\hat{W}(\hat{\mathbf{b}}_0)^{-1}$ as the weighting matrix in the second step where:[58]

$$\hat{W}(\hat{\mathbf{b}}_0) = \Big(\frac{1}{N}\sum_{i=1}^{N} f^{[b]}_{d_1,d_2}(\zeta^i;\hat{\mathbf{b}}_0) f^{[b]}_{d_1,d_2}(\zeta^i;\hat{\mathbf{b}}_0)^T\Big)^{-1}$$

There are several advantages of this approach. First, while Honoré (1993) provides continuous time identification results for duration models with multiple spells, the moment conditions used in the GMM estimator are based on discrete time identification results. Second, some approaches rely on specification of a frailty distribution. For example, Nakamura and Steinsson (2008) apply the empirical model of Meyer (1990) in their analysis of price spell duration and assume that the frailty parameter follows a gamma distribution for their baseline specification. Heckman and Singer (1984) note that misspecification of the frailty distribution can bias the hazard function. Instead, the approach of Alvarez et al. (2021) imposes no restrictions on the frailty distribution. Third, the GMM estimator is consistent when the number of individuals is large, but it allows for a short time dimension. The latter is important in our setting given that we rely on annual data covering 26 years.

[58] Hansen (1982) show that $\hat{W}(\hat{\mathbf{b}}_0)$ converges in probability to $\Omega \equiv \mathbb{E}[f^{[b]}_{d_1,d_2}(\zeta^i;\mathbf{b}_0) f^{[b]}_{d_1,d_2}(\zeta^i;\mathbf{b}_0)^T]$ and that $W = \Omega^{-1}$ is the most efficient weighting matrix.

D Alternative theories of participation

D.1 Nonstandard preferences

Expected utility maximizers with standard preferences exhibiting second-order risk aversion (e.g., CRRA utility) should always be willing to invest some money in stocks as long as the expected risk premium is positive (Haliassos and Bertaut, 1995). This is because such individuals are effectively risk neutral for small risks and risk has no first-order effect. As such, a model where all agents exhibit second-order risk aversion would need to be augmented with additional ingredients to motivate nonparticipation, such as participation costs or background risks. Some papers have allowed for time-varying levels of risk aversion, with one popular method being to have habit-formation preferences. Such preferences generate a negative relationship between wealth and risk aversion.[59] However, this will simply lead to time-varying risky asset shares with no impact on the extensive margin of participation as long as preferences still exhibit second-order risk aversion.[60]

To generate nonparticipation exclusively through preferences, first-order risk aversion is needed (Segal and Spivak, 1990).[61] Under such preferences, individuals have a kink in the utility function at some certainty point, which can make risk aversion locally infinite and zero stockholdings an optimal outcome. To generate dynamics in participation, we would need some agents to exhibit time-varying first-order risk aversion (Gomes et al., 2021). In addition, preferences would need to fluctuate at a reasonably high frequency to generate short-term dynamics. However, such models would likely struggle to explain the downward-sloping hazard functions for exit and reentry. Indeed, if such preference "shocks" have a constant Poisson arrival rate, the hazard rates should be flat. Furthermore, empirical studies have typically found positive and significant autocorrelations in individuals' risk preferences, suggesting that preferences are moderately stable, although correlations are usually below 1 (Chuang and Schechter, 2015; Dohmen et al., 2016).[62]

[59] These studies have typically used habit-formation preferences to help reproduce empirical patterns of equity premia (e.g., Constantinides, 1990; Campbell and Cochrane, 1999).

[60] Brunnermeier and Nagel (2008) empirically test whether wealth fluctuations affect risky asset shares and find no clear relationship, which they argue lends support to a CRRA model over a model with habit-formation preferences.

[61] A range of preferences exist that exhibit first-order risk aversion including, but not limited to prospect theory (Kahneman and Tversky, 1979), rank-dependent expected utility (Quiggin, 1982; Epstein and Zin, 1990), disappointment aversion (Gul, 1991; Ang et al., 2005; Routledge and Zin, 2010), news utility (Pagel, 2018) and ambiguity aversion (Gilboa and Schmeidler, 1989; Cao et al., 2005).

[62] Part of these imperfect correlations could reflect measurement error (Schildberg-Hörisch, 2018).

D.2 Risks faced by households

A strand of the literature studies how background risks, particularly labor income risk, can affect portfolio allocations. Theoretically, the impact of labor income risk depends on the nature of the risk (Vissing-Jørgensen, 2002). First, if labor income is riskless, this should lead to a higher investment in risky financial assets because such labor income is effectively equivalent to holding a riskless bond. Second, if labor income is risky but uncorrelated with stock returns, then individuals should tilt their portfolio away from stocks, as there is already risk coming from human wealth.[63] Third, if labor income is risky and correlated with stock returns, then there is a hedging component that runs with the opposite sign of the correlation. For example, if business cycle risk produces a positive correlation between labor income and stock returns, then the optimal portfolio choice requires one to reduce stockholdings (Haliassos and Bertaut, 1995). It is important to note that zero stockholding cannot be an optimal solution in the first two cases. Risky labor income that is uncorrelated with stock returns reduces the optimal portfolio share but would not push it to zero. However, Haliassos and Bertaut (1995) show that zero stockholding can be an optimal choice for sufficiently low wealth if labor income and stock returns are positively correlated, particularly if coupled with a no short-selling constraint. For a model to generate dynamics through labor income risk alone, we would require that 1) the correlation between labor income and stock returns is time-varying, and/or 2) wealth fluctuates around the participation threshold for some individuals, leading to entry and exit. The first can be hard to justify given that most individuals do not change jobs at a high frequency such that the underlying correlations could change. Regarding the second route, Figure 1.30b shows that short spells, while being relatively more likely for less wealthy individuals, still occur for high wealth groups at a nonnegligible frequency. In any case, empirical estimates for the correlation between labor income and stock returns are typically very close to zero, making such channels hard to rationalize from the data (e.g., Cocco et al., 2005; Fagereng et al., 2017a).

[63] Fagereng et al. (2017b) studies the impact of uninsurable wage risk on portfolio shares using Norwegian data. They find a significant marginal effect of such risk on portfolio shares, although the economic impact is limited because the size of this wage risk is small. Vissing-Jørgensen (2002) finds a negative impact of the volatility of nonfinancial income on both the probability of stock market participation and the proportion of wealth invested in stocks conditional on participating.

D.3 Cultural and social environment

Cultural factors can influence an individual's beliefs and preferences, which in turn affect economic outcomes (Guiso et al., 2006).[64] Various papers have provided empirical evidence of a causal link running from cultural environments to savings behavior.[65] While underparticipation in the stock market could be linked to cultural factors, these factors need to be time-varying to obtain dynamics in participation. However, Guiso et al. (2006) define culture as *"customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation"*. As such, cultural factors are very slow-moving and thus would not be able to reproduce the high frequency entry and exit that we observe.

However, social interactions could generate more frequent changes in beliefs and preferences. Shiller et al. (1984) argue that investing is a social activity, and therefore, investment decisions can be affected by the actions of those one interacts with. A growing literature has provided empirical evidence for the influence of peer effects on financial behavior.[66] In principle, communication between peers could lead to entry and exit. If my neighbor decides to leave the stock market – perhaps due to experiencing poor returns – this could induce me to also leave. If my neighbor claims that stock returns will be good in the near future, this might induce me back into the market. Testing these effects directly could be an interesting avenue for future research, although it seems unlikely that peer effects alone can explain all the dynamics we observe for a variety of reasons. First, Kaustia and Knüpfer (2012) show that good stock returns experienced by local peers can positively affect an individual's decision to enter the stock market. However, the authors do not find evidence of a discouragement effect following poor realizations, from which they infer that peers primarily share good outcomes with each other. Therefore, peer effects could struggle to explain exit. Second, it is difficult to rationalize the downward-sloping hazard functions through peers alone. Third, our focus is on the extensive margin of participation. We, therefore, require social interactions to generate complete exit rather than just exit from a particular stock. One could imagine individuals discussing partic-

[64] For example, ethnic origin has been shown to affect trust (Guiso et al., 2003b).

[65] Haliassos et al. (2017) study migrants to Sweden and find significant differences in financial behavior and the propensity to hold stocks based on the degree of cultural similarity to Sweden. Other papers that find significant effects of culture on financial behavior include Osili and Paulson (2008), Guin (2017), and Fuchs-Schündeln et al. (2020). However, some papers do not find such effects (Carroll et al., 1994; Carroll et al., 1999).

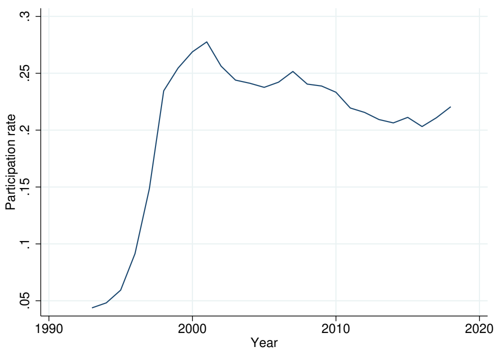
[66] Hong et al. (2004) show that households who report interacting with their neighbors and attending church are more likely to participate in the stock market even after controlling for individual characteristics and personality traits. Brown et al. (2008) find a causal link between individual stockholding and the average participation of the individual's community, which they argue occurs through word-of-mouth communication.

ular stocks, and perhaps a bad return experienced by a peer may deter them from also investing in that security. However, it may not necessarily put the person off investing in other stocks or funds.

E Additional tables and figures

Figure 1.22: Stock market participation rates over time by asset class

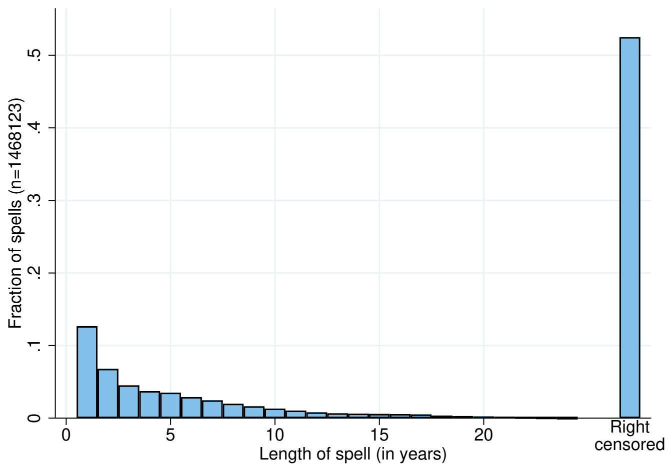
(a) Mutual funds



(b) Direct stockownership

Notes. This figure plots the participation rate in the stock market by asset class annually from 1993-2018. The left panel shows the participation rate in mutual funds, while the right panel is for directly held stocks.

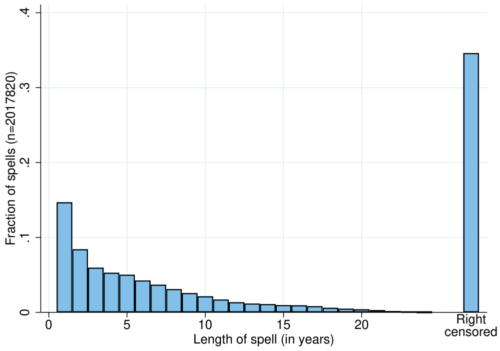
Figure 1.23: Spell length distribution at the household level



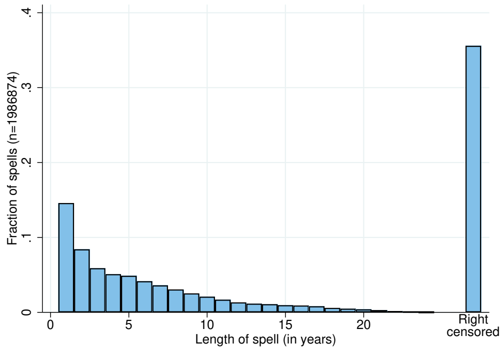
Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data based on the household-level balance sheet. A household is treated as participating in the stock market in year t if at least one spouse has some assets held in public equity. We take all spells beginning at any point from 1994-2015. The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.24: Spell length distribution (robustness to gifts/inheritance)

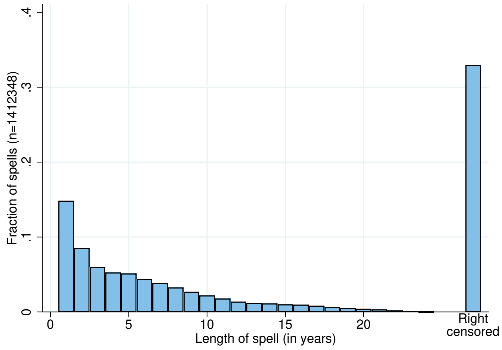
(A) No gift above 10,000 NOK



(B) No (grand)parent death

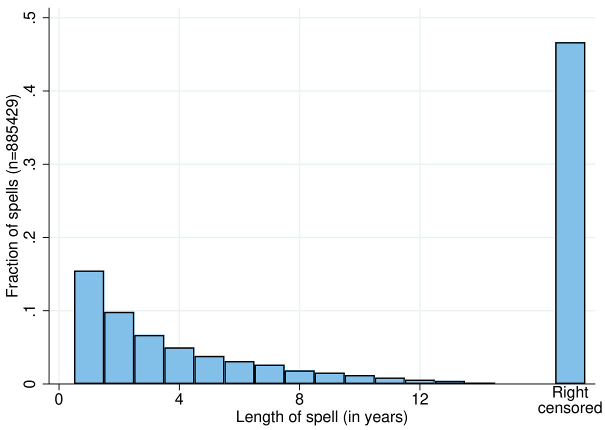


(C) No (grand)parent participation



Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data for different subsamples intended to deal with concerns that short spells are driven by gifts and inheritances. For all panels, we take spells beginning at any point from 1994-2015. The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored. Panel (A) excludes all individuals who receive a gift or inheritance above 10,000 NOK (based on tax records) in the year of or before entry. Panel (B) excludes all entrants who experience the death of a parent or grandparent in the year of or before entry. Panel (C) excludes all entrants for whom a parent or grandparent participated in the year of or before entry.

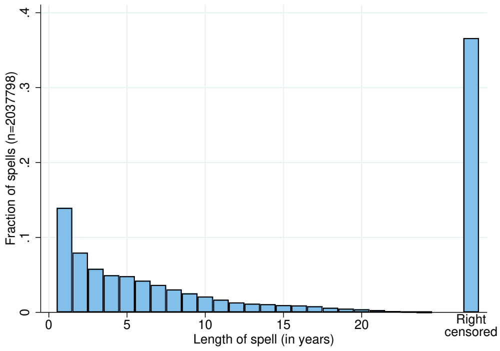
Figure 1.25: Spell length distribution excluding employee stocks



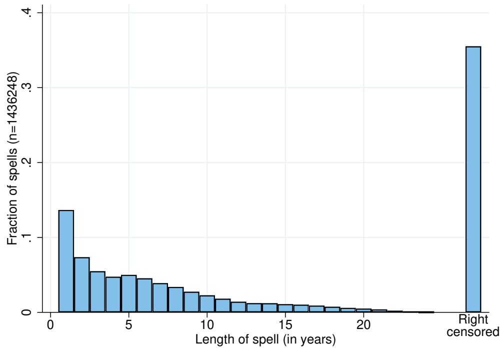
Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data excluding entrants who hold stocks in the company they work for. Such individuals are identified using the Shareholder Registry and demographic information about place of work. We take all spells beginning at any point from 2004-2015 (Shareholder registry data begin in 2004). The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.26: Spell length distribution excluding small investors

(a) Invest > $100

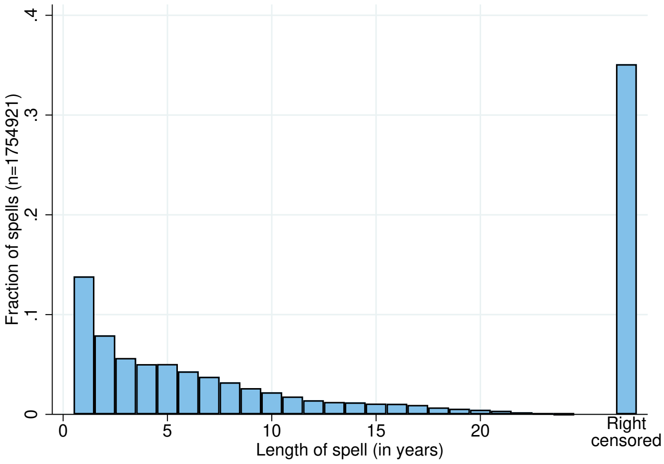


(b) Invest > $1000



Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data excluding entrants who invest a "small" amount of money at the point of entry. The left panel only uses individuals who invest at least $100 at the point of entry, while the right panel requires an investment of at least $1,000. For both panels, we take spells beginning at any point from 1994-2015. The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.27: Spell length distribution using only first spells

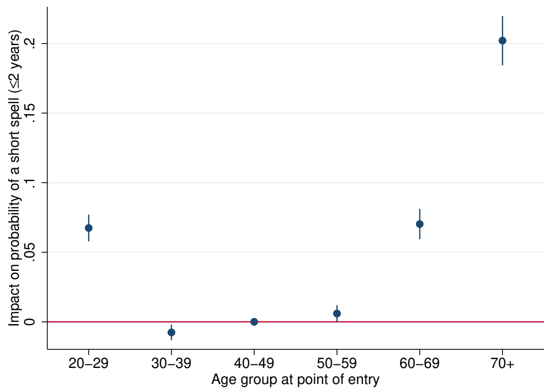


Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data using only the first recorded spell of a given participant. We take all first spells beginning at any point from 1994-2015. The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.28: Wealth distribution by spell length

Notes. This figure plots the proportion of stock market participants belonging to different wealth deciles as measured at the point of entry, separately for short spellers (participate for ≤ 2 years) and longer-term participants (> 2 years).

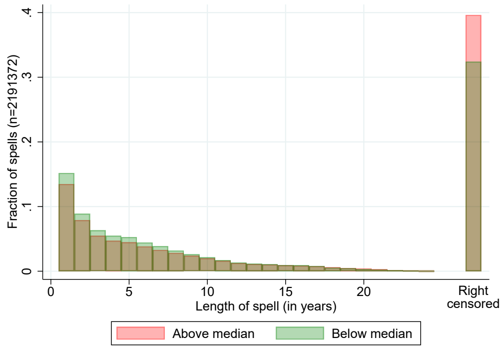
Figure 1.29: Impact of age on the probability of a short spell



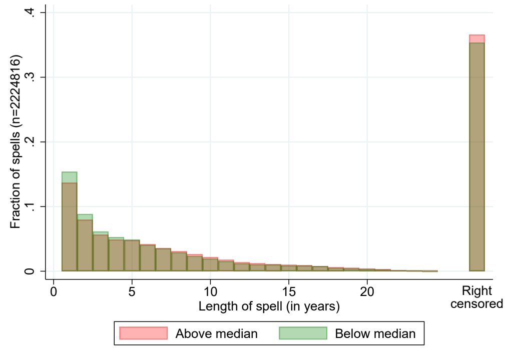
Notes. This figure plots the coefficient estimates for the fixed effects on age following estimation of Equation 1.1. Individual fixed effects are included in this specification. Age is measured at the point of entry and individuals are grouped into 10-year bins. 95% confidence intervals are shown. The red line represents a null relative effect.

Figure 1.30: Spell length distribution by observable characteristics

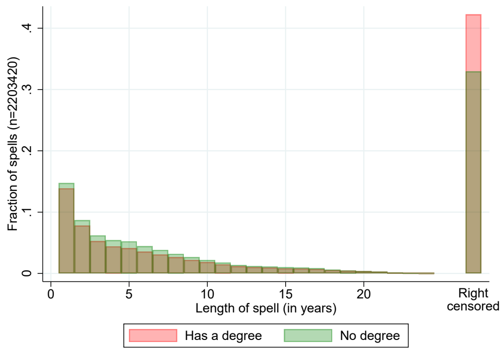
(A) Income



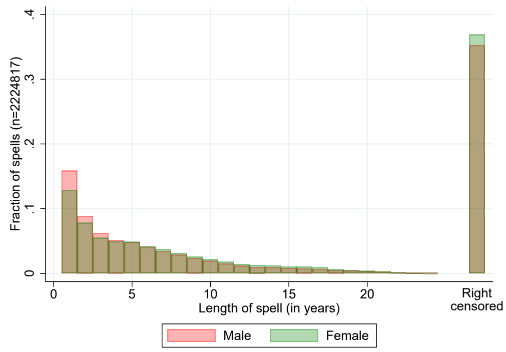
(B) Wealth



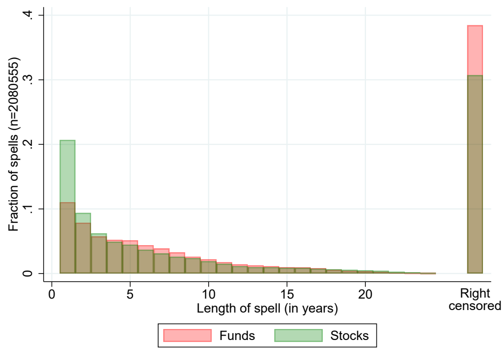
(C) Education



(D) Gender



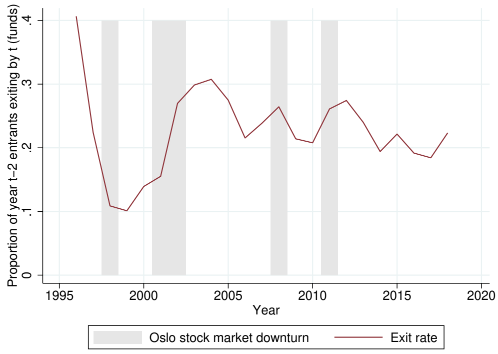
(E) By asset class



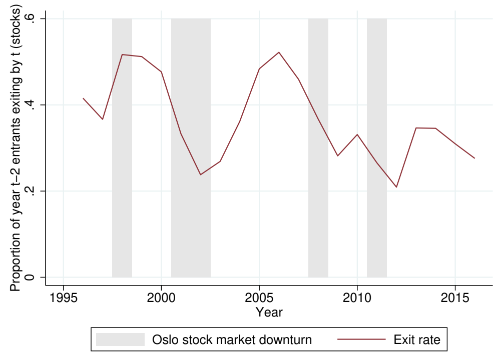
Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data for different observable characteristics, namely income (A), wealth (B), education (C), and gender (D). Panel (E) looks at individuals who enter into mutual funds vs. directly held stocks. For this panel, we exclude those entrants who choose to invest in both at the point of entry. For all panels, we take all spells beginning at any point from 1994-2015. The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.31: Prevalence of short spells over time by asset class

(A) Mutual funds



(B) Direct stockholding



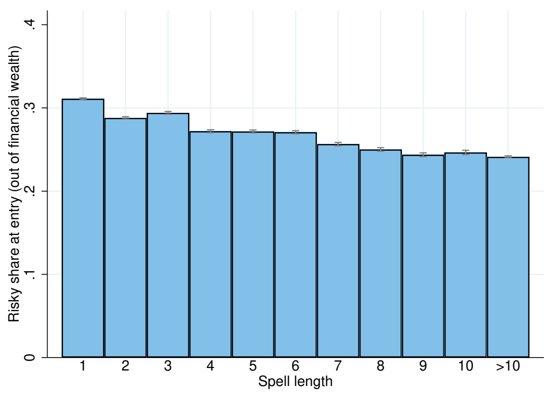
Notes. This figure plots the proportion of year $t-2$ entrants who leave the stock market by year t separately based on asset class. Panel (A) corresponds to individuals who entered into mutual funds, while panel (B) is for direct stockholders. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

Table 1.4: Summary statistics by type of individual

	Never	Always	Single short	Multiple
Male	0.43	0.51	0.50	0.60
Single	0.38	0.31	0.35	0.31
College degree	0.20	0.34	0.25	0.38
Unemployed	0.06	0.05	0.07	0.05
Financial wealth (2011 $'000s)	26.53	74.88	34.63	71.66
Total wealth (2011 $'000s)	153.43	281.51	199.63	300.67
Income (2011 $'000s)	32.29	50.18	43.66	60.10

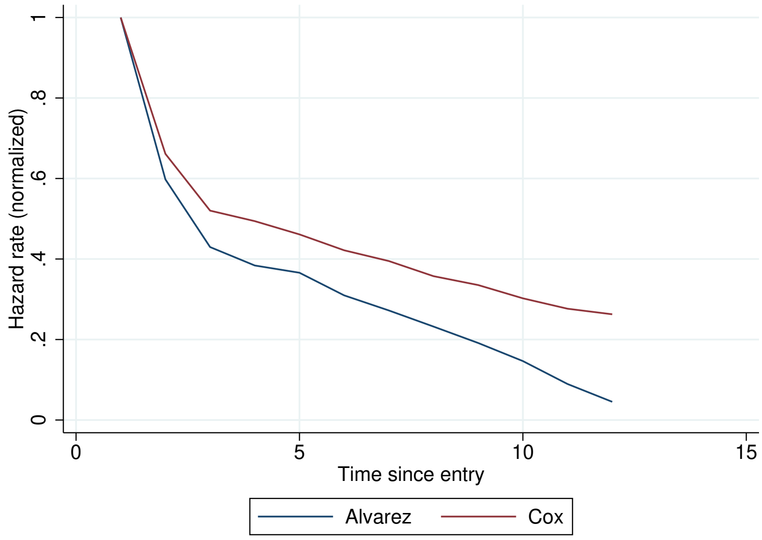
Notes. This table provides summary statistics of a range of characteristics for 4 groups: individuals who never participate in the stock market ("Never"), those who have one single spell lasting at least 5 years ("Always"), those who have one spell lasting less than 5 years ("Single short"), and individuals who have at least two spells in the stock market ("Multiple"). Mean values are reported.

Figure 1.32: Share of financial wealth invested in public equity at point of entry by spell length



Notes. This figure plots the average risky share (amount invested in public equity out of total financial wealth) at the point of entry separately for individuals of different eventual spell lengths. 95% confidence intervals are shown.

Figure 1.33: Cox proportional hazard function for exit from participation



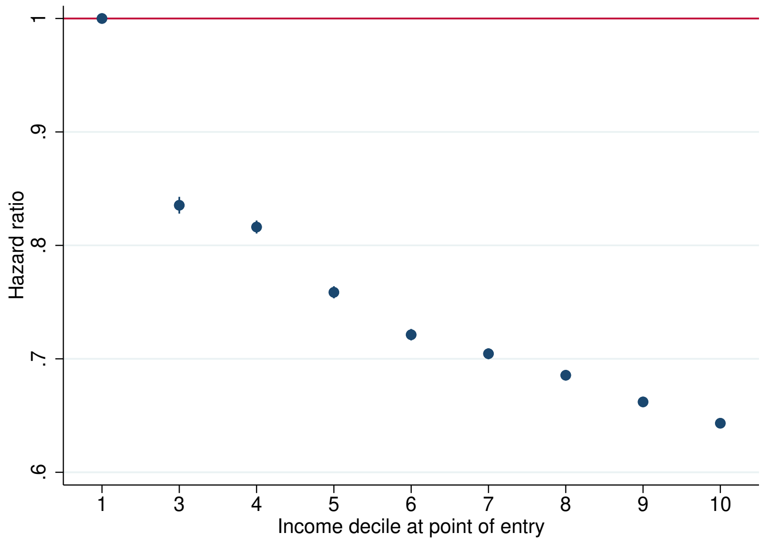
Notes. This figure plots the baseline hazard for exit from participation estimated using a Cox proportional hazards model (see Equation 1.2). The baseline hazard estimated using the Alvarez et al. (2021) methodology (see Section 3.1.2) is also shown. To faciliate comparison with the Alvarez et al. (2021) hazard function, the Cox baseline hazard has been normalized to 1 at duration $d = 1$. Hazard ratios (exponentiated coefficients) for the covariates X_i are given in Table 1.5 and Figures 1.34-1.36.

Table 1.5: Hazard ratios from Cox proportional hazards estimation (exit from participation)

	Hazard ratio
Male	1.214***
	(106.63)
College degree	0.923***
	(-42.06)
Single	1.144***
	(73.70)
Individual FE	No
Entry year FE	Yes
Age group FE	Yes
Income decile FEs	Yes
Wealth decile FEs	Yes
Observations	18076414

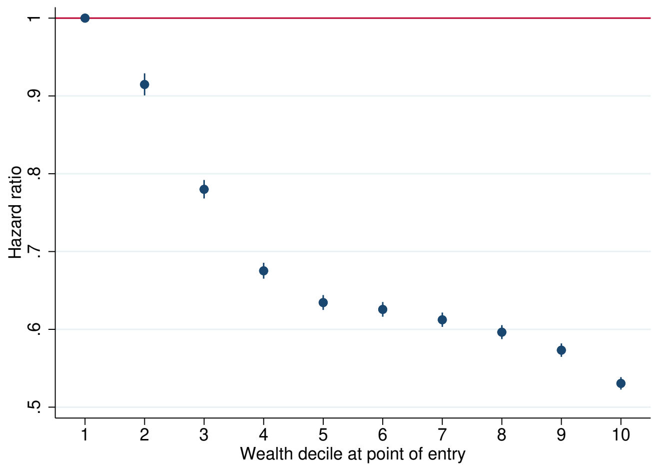
Notes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. This table shows hazard ratios from estimation of a Cox proportional hazards model for exit from participation (Equation 1.2). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for exiting from the stock market associated with a one-unit increase in the covariate.

Figure 1.34: Hazard ratios on income deciles in Cox model (exit from participation)



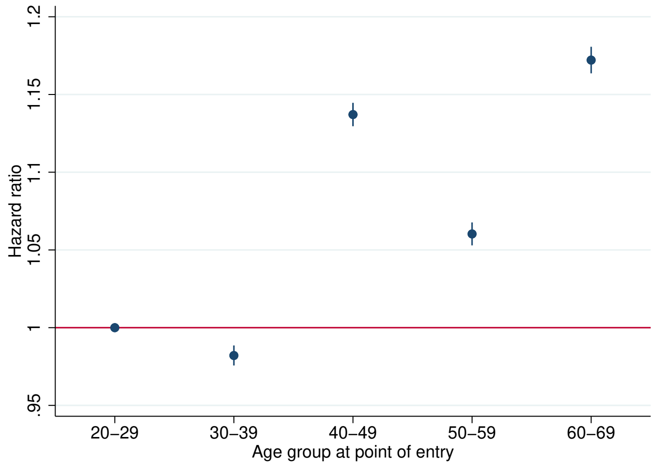
Notes. This figure plots the hazard ratios on income deciles in the Cox proportional hazards model for exit from participation (see Equation 1.2). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for exiting from the stock market associated with a one-unit increase in the covariate (relative to the first decile, which is omitted to avoid collinearity).

Figure 1.35: Hazard ratios on wealth deciles in Cox model (exit from participation)



Notes. This figure plots the hazard ratios on wealth deciles in the Cox proportional hazards model for exit from participation (see Equation 1.2). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for exiting from the stock market associated with a one-unit increase in the covariate (relative to the first decile, which is omitted to avoid collinearity).

FIGURE 1.36: Hazard ratios on age in Cox model (exit from participation)



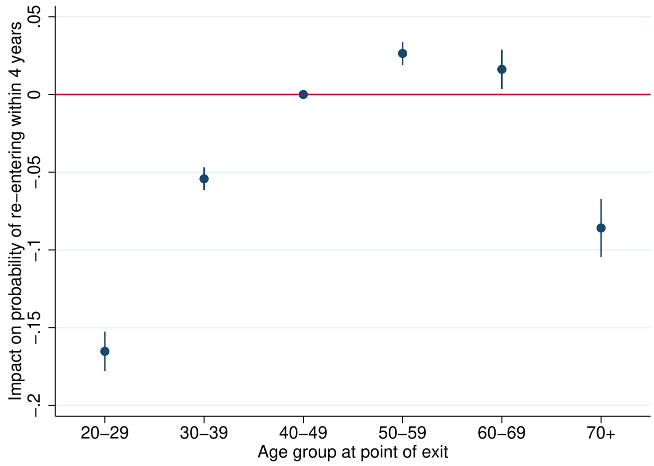
Notes. This figure plots the hazard ratios on age in the Cox proportional hazards model for exit from participation (see Equation 1.2). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for exiting from the stock market associated with a one-unit increase in the covariate (relative to the youngest age group, which is omitted to avoid collinearity).

TABLE 1.6: Determinants of reentry

	Reentry in 4y	
Male	0.037***	
	(0.001)	
College degree	0.030***	0.011
	(0.001)	(0.007)
Homeowner	-0.063***	-0.022***
	(0.001)	(0.004)
Unemployed	-0.005**	0.001
	(0.002)	(0.004)
Single	-0.028***	-0.056***
	(0.001)	(0.003)
Sample mean	0.35	0.59
Individual FE	No	Yes
Exit-year FE	Yes	Yes
Age group FE	Yes	Yes
Income decile FE	Yes	Yes
Wealth decile FE	Yes	Yes
Observations	1436019	518995
R-squared	0.14	0.54

Notes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. This table shows the estimation of the linear probability model in Equation 1.3. The first column excludes individual fixed effects, while the second column includes them. The dependent variable is a binary variable equal to 1 if the exiter reenters within 4 years following exit, and zero otherwise. *Homeowner* equals 1 if the participant owns their own property (either self-owned or ownership through housing cooperatives), and zero otherwise. *Single* equals 1 if the participant is neither married nor cohabiting, and zero otherwise. *Unemployed* equals 1 if the participant receives unemployment benefits at the point of exit, and zero otherwise. Exit-year fixed effects are included. Age fixed effects by broad age group (20-29, 30-39, 40-49, 50-59, 60-69 and 70+), as well as income and wealth decile fixed effects are included. Observables are measured at the point of exit. Standard errors are clustered at the individual level. The regression uses data on exiters from 1994-2014.

FIGURE 1.37: Impact of age on the probability of reentry



Notes. This figure plots the coefficient estimates for the age group fixed effects following estimation of Equation 1.3. This specification includes individual fixed effects. Age is measured at the point of exit, and individuals are grouped into 10-year bins. 95% confidence intervals are shown. The red line represents a null relative effect.

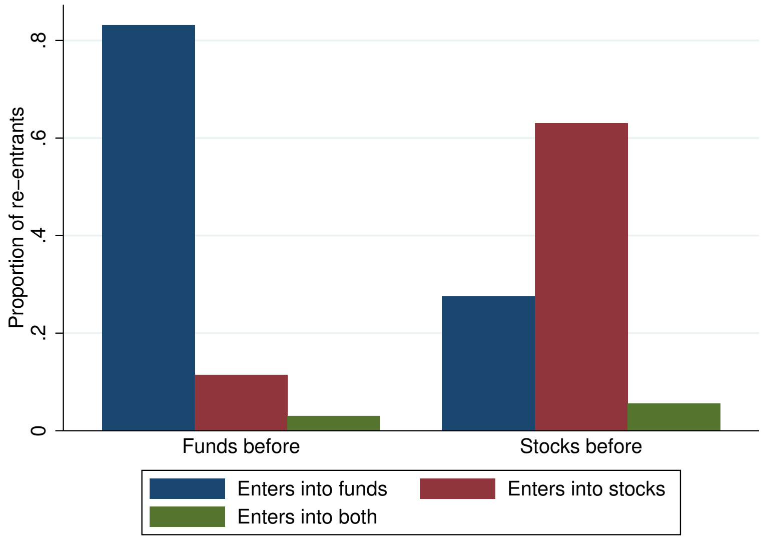
FIGURE 1.38: Exit rate over time by asset class

(A) Mutual funds

(B) Direct stockholding

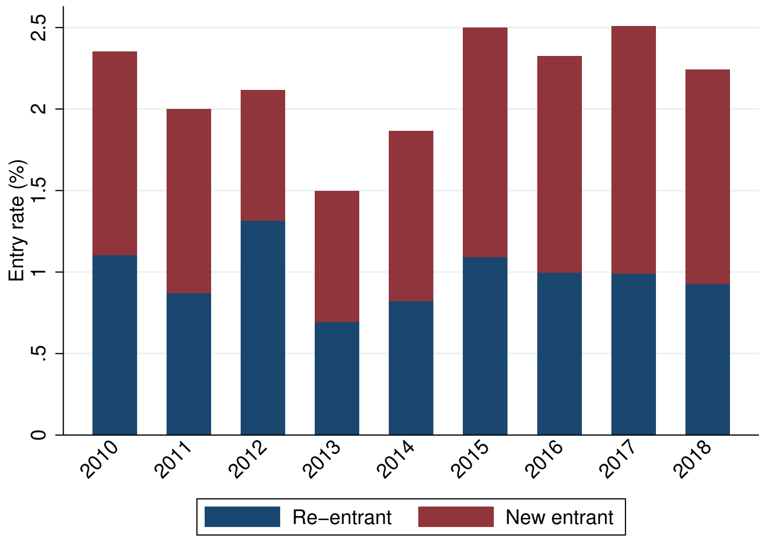
Notes. This figure plots the exit rate separately for mutual funds (A) and direct stockholding (B). The exit rate in year t is computed as the proportion of participants in year $t-1$ who leave that asset class in year t. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

Figure 1.39: Reentry into different asset classes by previous asset class choice



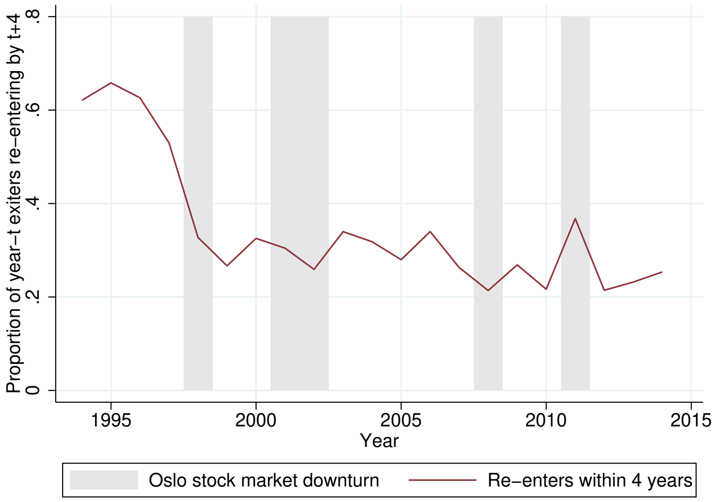
Notes. This figure plots the proportion of reentrants going into funds, stocks or both by the choice of funds vs. stocks in their previous spell.

Figure 1.40: Decomposition of entry rate into reentrants and new entrants



Notes. This figure decomposes the stock market entry rate in a given year into two components: entry by former participants ("Re-entrant") and new entrants who have not participated before. The entry rate in year t is the proportion of nonparticipants in year $t-1$ who enter in year t.

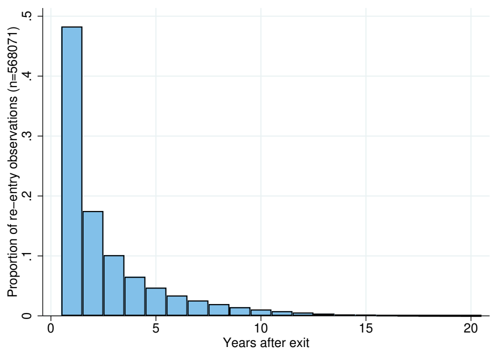
Figure 1.41: Reentry rate over time



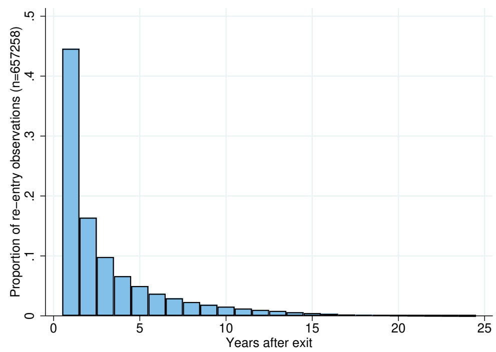
Notes. This figure plots the proportion of exiters of a given year who reenter within the next 4 years. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

Figure 1.42: Distribution of reentry times (robustness to gifts/inheritance)

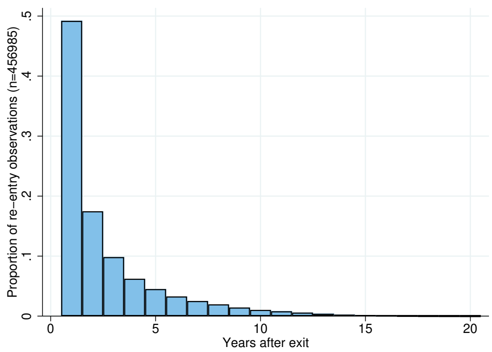
(A) No gift above 10,000 NOK



(B) No (grand)parent death

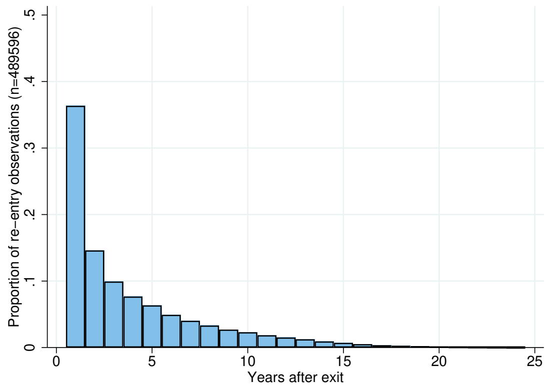


(C) No (grand)parent participation



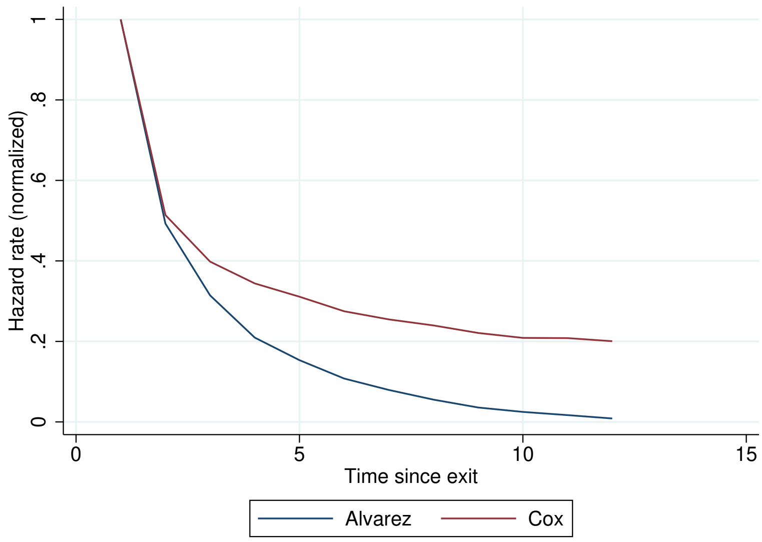
Notes. This histogram shows the distribution of reentry times in the Norwegian data for different subsamples intended to deal with concerns that short spells are driven by gifts and inheritances. The x-axis gives the reentry time (in years) and the y-axis shows the proportion of reentry observations belonging to a particular length. Panel (A) excludes all reentrants who receive a gift or inheritance above 10,000 NOK (based on tax records) in the year of or before reentry. Panel (B) excludes all reentrants who experience the death of a parent or grandparent in the year of or before reentry. Panel (C) excludes all reentrants for whom a parent or grandparent was participating in the year of or before reentry.

Figure 1.43: Distribution of reentry times (excluding employee stocks)



Notes. This histogram shows the distribution of reentry times in the Norwegian data excluding reentrants who hold stocks in the company they work for. The x-axis gives the reentry time (in years) and the y-axis shows the proportion of reentry observations belonging to a particular length. As the Shareholder Registry data are only available from 2004, we only consider reentry observations where the year of reentry is no earlier than 2004.

Figure 1.44: Cox proportional hazard function for reentry



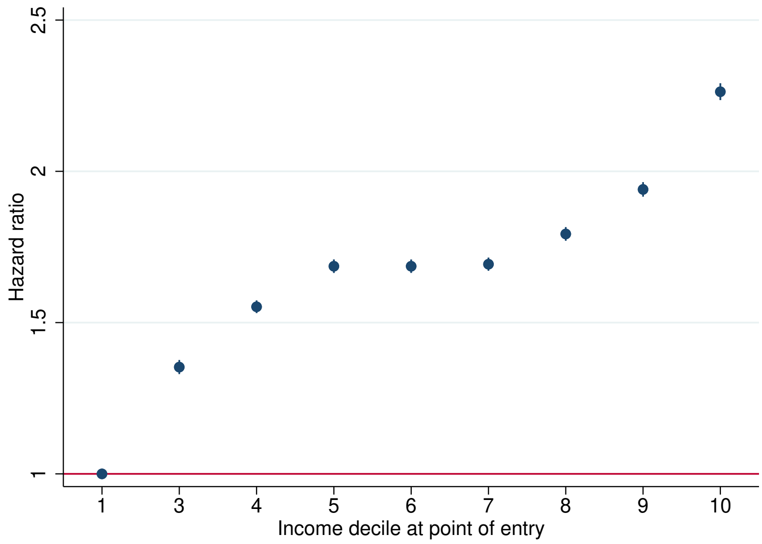
Notes. This figure plots the baseline hazard for reentry following exit estimated using a Cox proportional hazards model (see Equation 1.4). The baseline hazard estimated using the Alvarez et al. (2021) methodology (see Section 3.2.3) is also shown. To faciliate comparison with the Alvarez et al. (2021) hazard function, the Cox baseline hazard has been normalized to 1 at duration $d = 1$. Hazard ratios (exponentiated coefficients) for the covariates X_i are given in Table 1.7 and Figures 1.45-1.47.

TABLE 1.7: Hazard ratios from Cox proportional hazards estimation (reentry)

	Hazard ratio
Male	1.180***
	(56.54)
College degree	1.133***
	(45.10)
Single	0.923***
	(-26.89)
Individual FE	No
Entry year FE	Yes
Age group FE	Yes
Income & wealth decile FEs	Yes
Observations	9949687

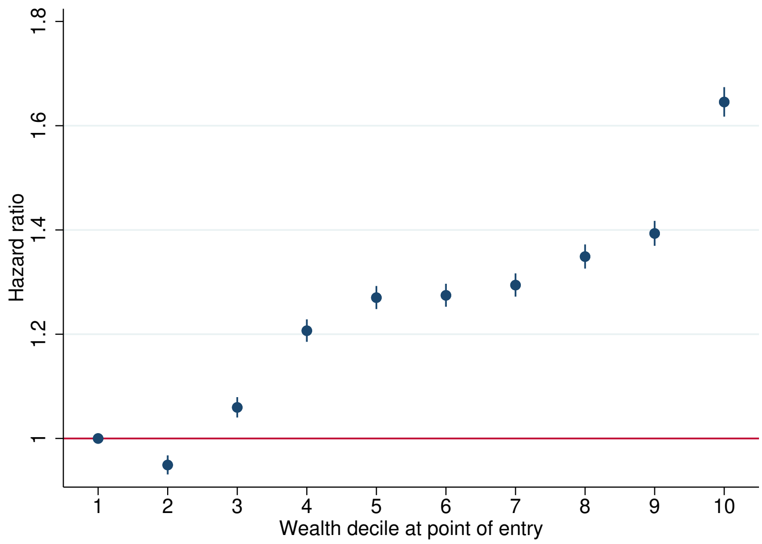
Notes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. This table shows hazard ratios from estimation of a Cox proportional hazards model for reentry following exit (Equation 1.4). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for exiting from the stock market associated with a one-unit increase in the covariate.

FIGURE 1.45: Hazard ratios on income deciles in Cox model (reentry)



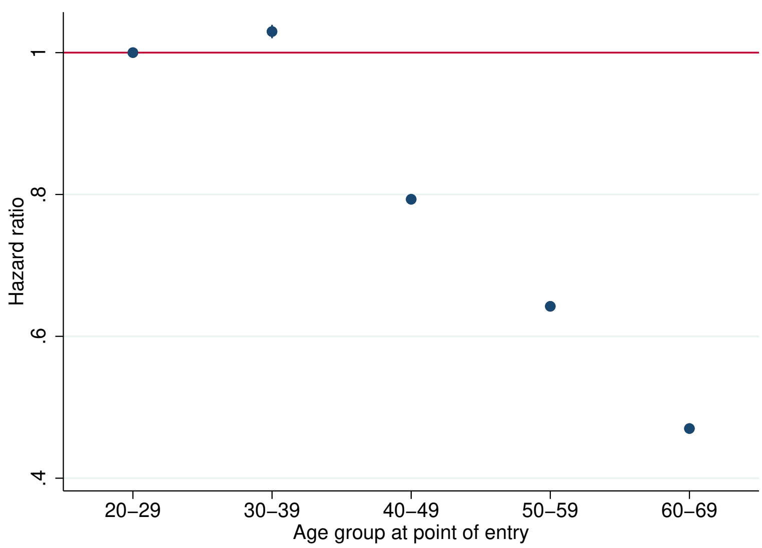
Notes. This figure plots the hazard ratios on income deciles in the Cox proportional hazards model for reentry following exit (see Equation 1.4). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for reentering the stock market associated with a one-unit increase in the covariate (relative to the first decile, which is omitted to avoid collinearity).

FIGURE 1.46: Hazard ratios on wealth deciles in Cox model (reentry)



Notes. This figure plots the hazard ratios on wealth deciles in the Cox proportional hazards model for reentry following exit (see Equation 1.4). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for reentering the stock market associated with a one-unit increase in the covariate (relative to the first decile, which is omitted to avoid collinearity).

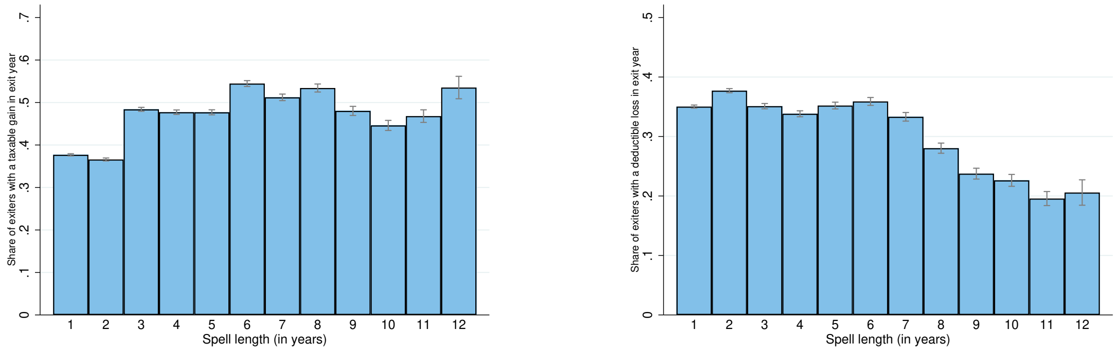
FIGURE 1.47: Hazard ratios on age in Cox model (reentry)



Notes. This figure plots the hazard ratios on age in the Cox proportional hazards model for reentry following exit (see Equation 1.4). Hazard ratios are exponentiated coefficients ($\exp(\hat{\beta})$), where $\hat{\beta}$ denotes the estimated coefficients. The hazard ratio gives the relative increase or decrease in the hazard for reentering the stock market associated with a one-unit increase in the covariate (relative to the youngest age group, which is omitted to avoid collinearity).

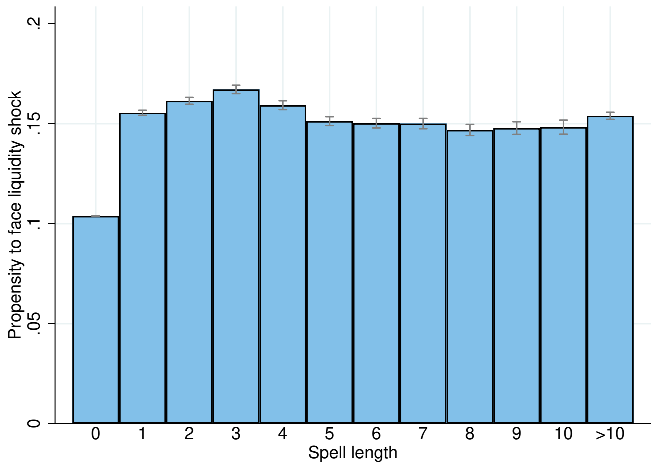
Figure 1.48: Performance of exiters by spell length

(A) Report gains (B) Report losses



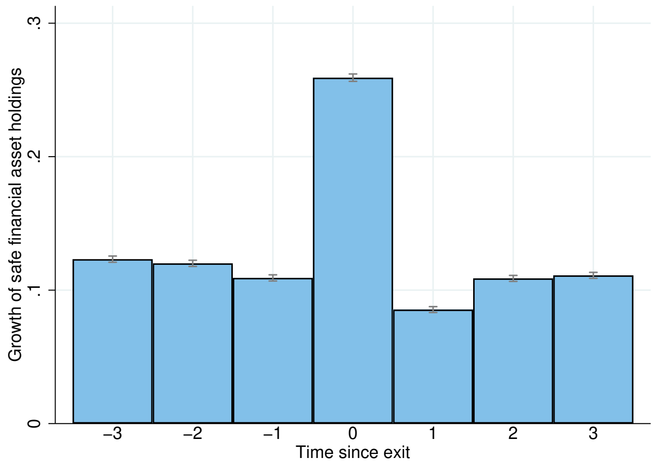
Notes. This figure shows the performance of exiters by spell length based on records of taxable gains and tax-deductible losses in the income tax data. In panel (A), we plot the proportion of exiters of a given spell length reporting some gains (irrespective of losses) from the sale of stocks and funds (gains are computed as the sum of items TR 3.1.8, TR 3.1.9, and TR 3.1.10 in the tax records) in their exit year. In panel (B), we plot the proportion of exiters reporting some losses (irrespective of gains). Losses are computed as the sum of items TR 3.3.8, TR 3.3.9, and TR 3.3.10 in the tax records. We use exiters who enter from 2006 onward in these plots.

Figure 1.49: Prevalence of liquidity shocks in exit year by spell length



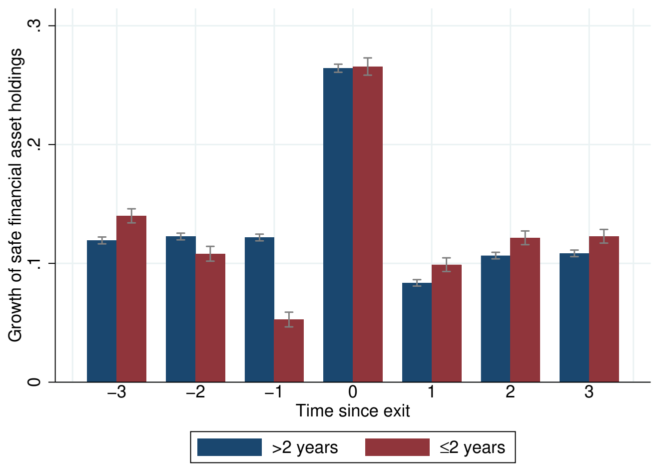
Notes. This figure shows the proportion of exiters of different spell lengths experiencing at least one of four potential liquidity needs in the exit year. The four shocks considered are buying a house (observed in housing transactions data), divorce, unemployment (inferred through receipt of unemployment benefits), and a large fall in income of $> 50\%$. The far-left bar (spell length of zero) gives the prevalence of liquidity shocks over nonexit observations (i.e. continuing participants). The far-right bar groups all exiters with spell lengths above 10 years. 95% confidence intervals are shown.

Figure 1.50: Average safe financial asset growth around exit year



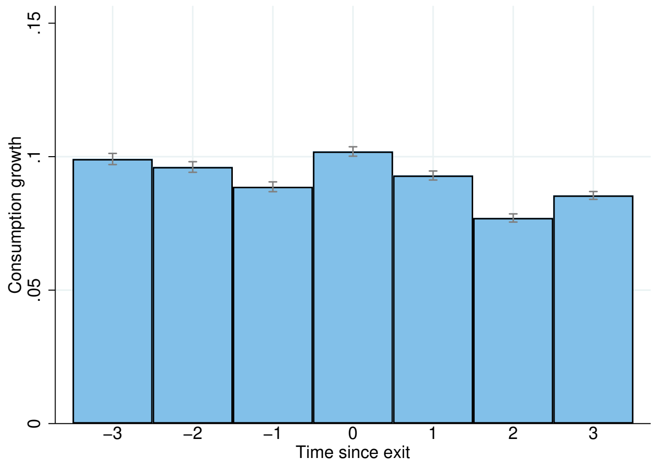
Notes. This figure plots the average growth rate of safe financial asset holdings in the year of exit from the stock market, as well as years either side of exit. Safe financial assets consists of cash, deposits and money market/bond funds. Growth rates are trimmed at the 5^{th} and 95^{th} percentiles. The analysis is based on household-level safe financial asset holdings, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.51: Average safe financial asset growth around exit year by spell length



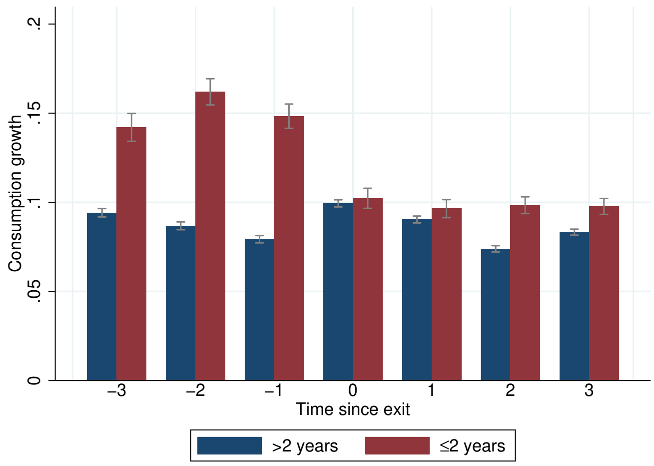
Notes. This figure plots the average growth rate of safe financial asset holdings separately for short spellers (exit within 2 years) and longer-term participants in the year of exit from the stock market, as well as years either side of exit. Safe financial assets consists of cash, deposits and money market/bond funds. Growth rates are trimmed at the 5^{th} and 95^{th} percentiles. The analysis is based on household-level safe financial asset holdings, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.52: Average consumption growth around exit year



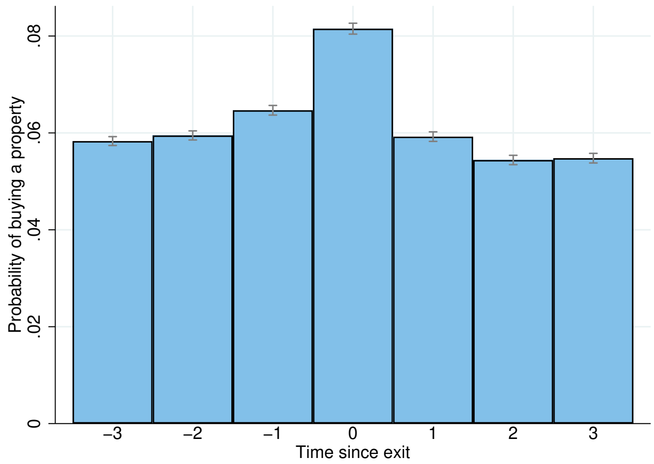
Notes. This figure plots the average consumption growth rate in the year of exit from the stock market, as well as years either side of exit. Growth rates are trimmed at the 5th and 95th percentiles. The analysis is based on household-level consumption, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown. Consumption data is based on debit card transactions data (see Aastveit et al., 2020).

Figure 1.53: Average consumption growth around exit year by spell length



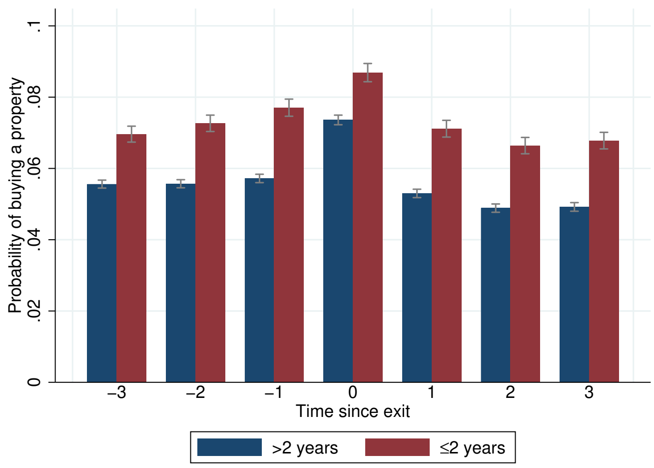
Notes. This figure plots the average consumption growth rate separately for short spellers (exit within 2 years) and longer-term participants in the year of exit from the stock market, as well as years either side of exit. Growth rates are trimmed at the 5th and 95th percentiles. The analysis is based on household-level consumption, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown. Consumption data is based on debit card transactions data (see Aastveit et al., 2020).

Figure 1.54: Change in house purchases around exit year



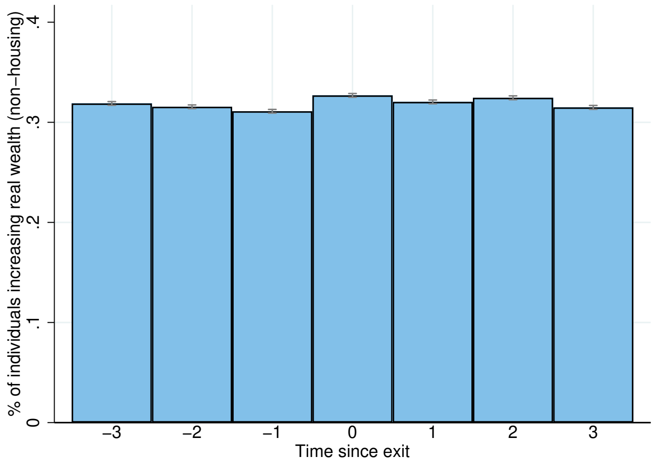
Notes. This figure plots the proportion of people buying a house in the year of exit from the stock market, as well as years either side of exit. The analysis is based on household-level homeownership, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.55: Change in house purchases around exit year by spell length



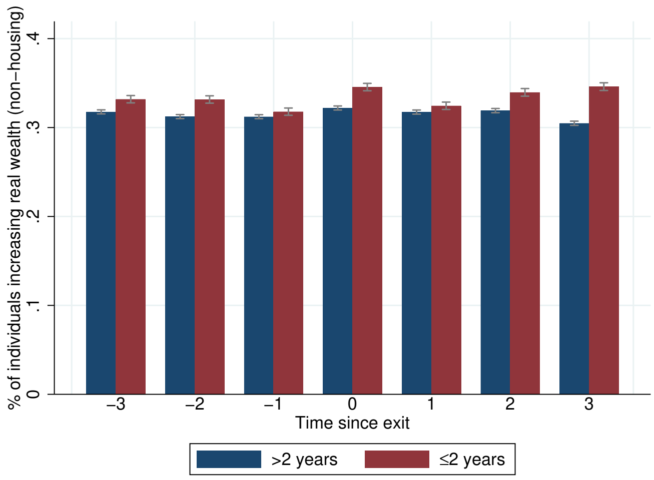
Notes. This figure plots the proportion of people buying a house separately for short spellers (exit within 2 years) and longer-term participants in the year of exit from the stock market, as well as years either side of exit. The analysis is based on household-level homeownership, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.56: Change in nonhousing real assets around exit year



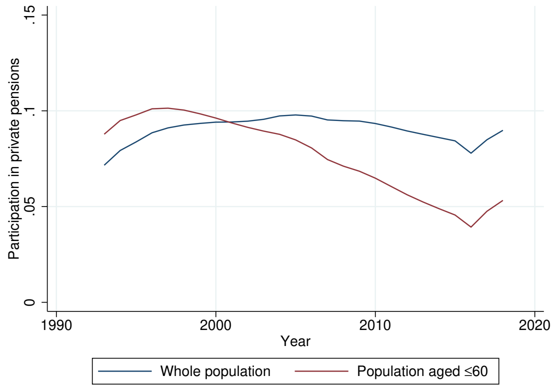
Notes. This figure plots the proportion of people increasing their nonhousing real asset holdings in the year of exit from the stock market, as well as years either side of exit. The analysis is based on household-level holdings, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.57: Change in nonhousing real assets around exit year by spell length



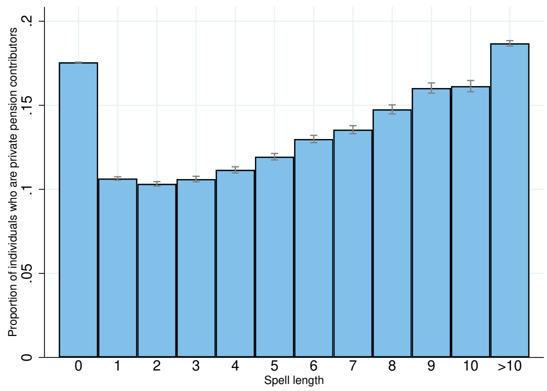
Notes. This figure plots the proportion of people increasing their nonhousing real asset holdings in the year of exit from the stock market, as well as years either side of exit. The analysis is based on household-level holdings, and uses only those households with at least $5,000 in the stock market in the year before exit. 95% confidence intervals are shown.

Figure 1.58: Participation in private pensions over time



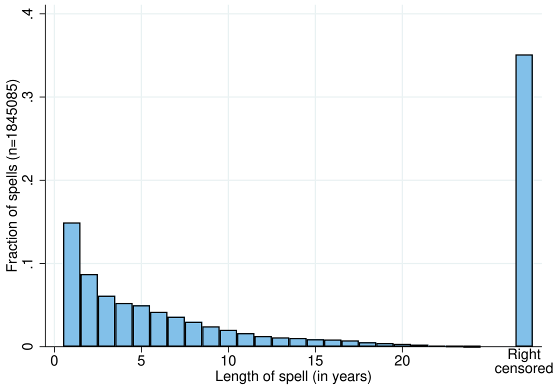
Notes. This figure plots a time series of participation in private pensions over time. The blue line gives the participation rate for the whole population, while the red line restricts attention to those aged 60 or under. An individual is said to be participating in private pensions in a given year t if they put money into a private pension either in the current year or in a past year. Participation has occurred if either of the following two items in the tax records is nonzero: TR 3.3.5, which records deductible payments to an Individual Pension Scheme (IPS), or TR 4.5.1, which records capital in an Individual Pension Account (IPA).

Figure 1.59: Prevalence of private pensions amongst exiters by spell length



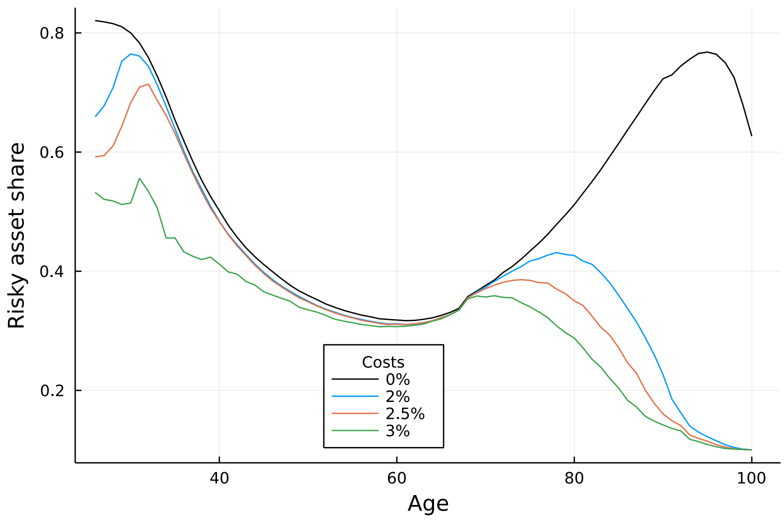
Notes. This figure shows the proportion of exiters of different spell lengths participating in private pension accounts as of their exit year. An individual is said to be participating in private pensions in their exit year if they put money into a private pension either in the current year or in a past year. Participation has occurred if either of the following two items in the tax records is nonzero: TR 3.3.5, which records deductible payments to an Individual Pension Scheme (IPS), or TR 4.5.1, which records capital in an Individual Pension Account (IPA). The far-left bar (spell length of zero) gives the prevalence of private pensions shocks over nonexit observations (i.e. continuing participants). The far-right bar groups all exiters with spell lengths above 10 years. 95% confidence intervals are shown.

Figure 1.60: Spell length distribution excluding individuals with a private pension account



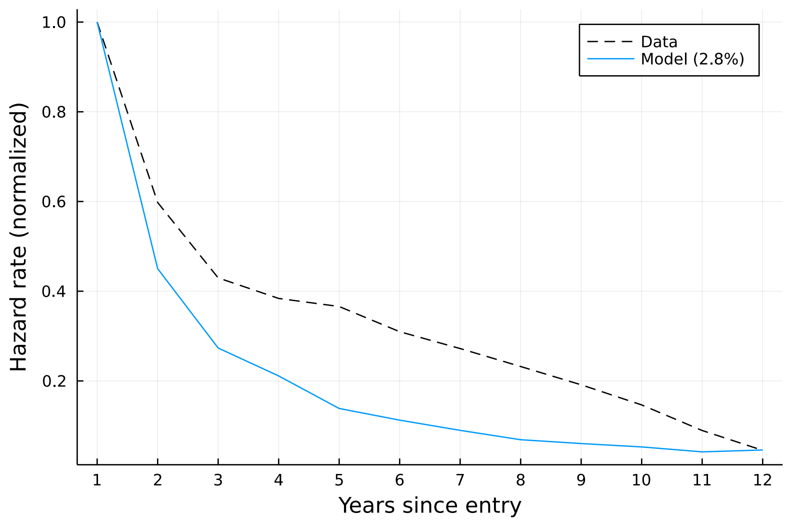
Notes. This histogram plots the proportion of spells of different lengths in the Norwegian data excluding individuals who at any point in the sample hold a private pension account. Participation has occurred if either of the following two items in the tax records is nonzero: TR 3.3.5, which records deductible payments to an Individual Pension Scheme (IPS), or TR 4.5.1, which gives capital in an Individual Pension Account (IPA). The x-axis gives the spell length (in years) and the y-axis shows the proportion of spells belonging to a particular spell length. The far-right bar gives the proportion of these spells that are right censored.

Figure 1.61: Model without beliefs: conditional risky share



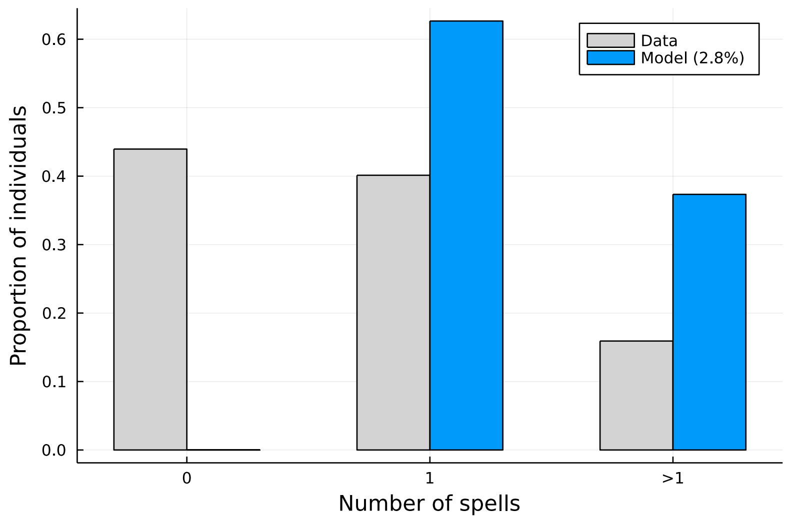
Notes. This figure plots the average risky asset share α_{it} conditional on participating in the stock market ($\alpha_{it} > 0$) for the model without beliefs ($b_{it} = 1\ \forall i, t$). The share is plotted for different values of per-period costs. Entry costs are set to zero in all cases.

Figure 1.62: Model without beliefs: hazard rate for exit



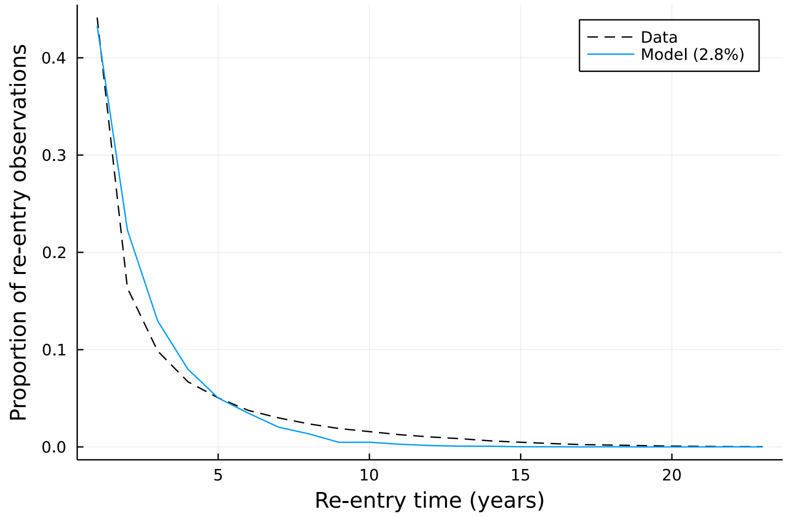
Notes. This figure plots the hazard rate for exit in the model without beliefs ($b_{it} = 1\ \forall i, t$). Per-period costs $\bar{F}^1$ are set at 2.8% of permanent income and entry costs are set to zero. The hazard rate at 1 year after entry is normalized to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified (see Section 3.1.2).

Figure 1.63: Model without beliefs: number of spells



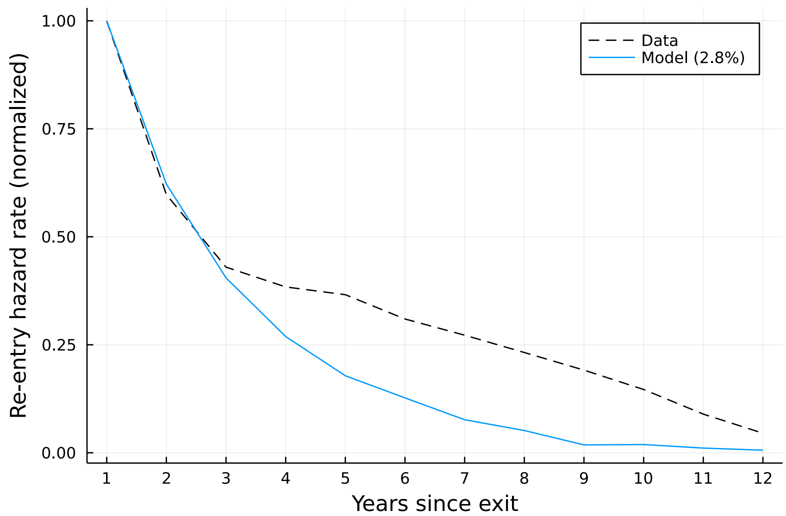
Notes. This figure plots the distribution of the number of spells in the simulated population for the model without beliefs ($b_{it} = 1\ \forall i, t$). Per-period costs $\bar{F}^1$ are set at 2.8% of permanent income and entry costs are set to zero. The empirical distribution for the Norwegian population is also shown (see Figure 1.9).

Figure 1.64: Model without beliefs: reentry times



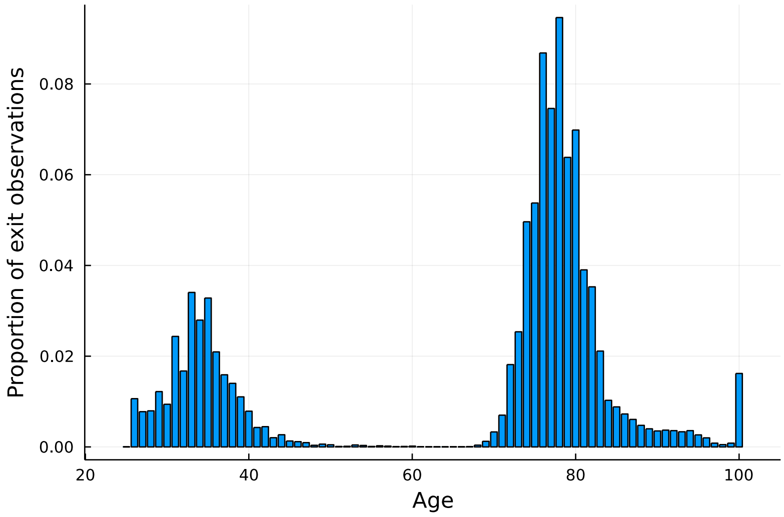
Notes. This figure plots the distribution of reentry times in the model without beliefs ($b_{it} = 1 \ \forall i, t$). Per-period costs $\bar{F}^1$ are set at 2.8% of permanent income and entry costs are set to zero. The empirical proportion from the Norwegian data is also shown (see Figure 1.11).

Figure 1.65: Model without beliefs: hazard rate for reentry



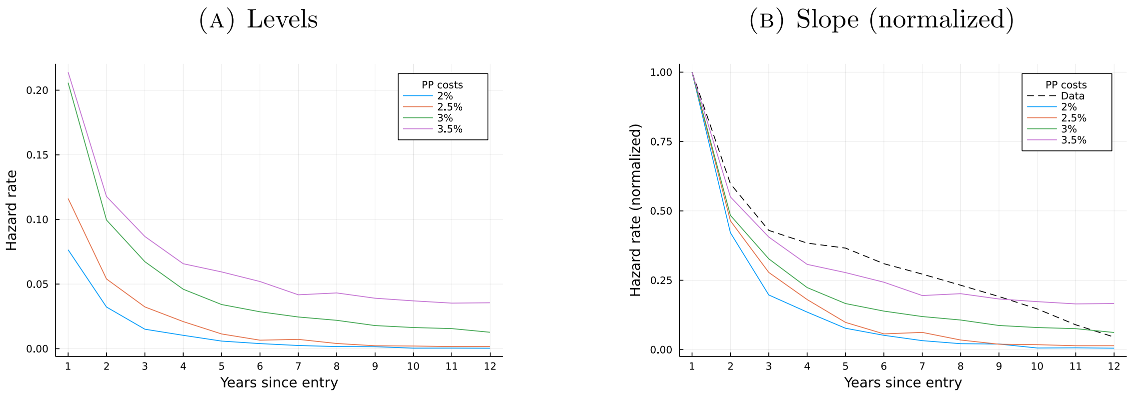
Notes. This figure plots the hazard rate for reentry in the model without beliefs ($b_{it} = 1 \ \forall i, t$). Per-period costs $\bar{F}^1$ are set at 2.8% of permanent income and entry costs are set to zero. The hazard rate at 1 year after exit is normalized to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified (see Sections 3.1.2 and 3.2.3).

Figure 1.66: Model without beliefs: exit points by age



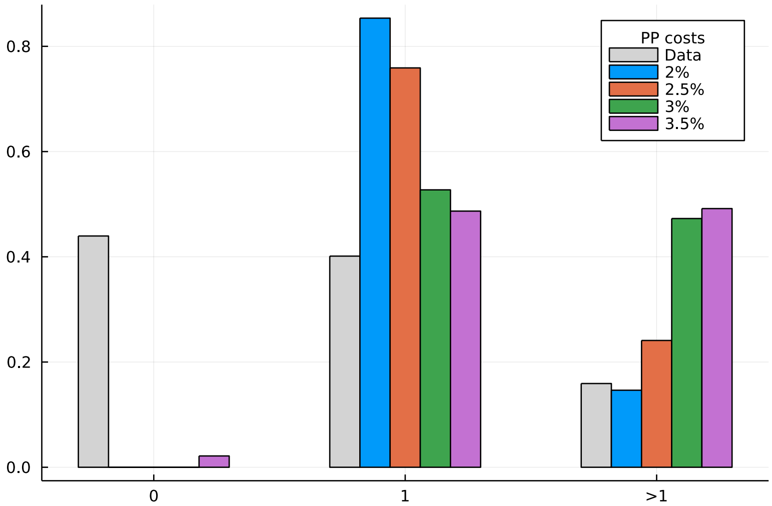
Notes. This figure plots the proportion of exit observations by age in the model without beliefs ($b_{it} = 1\ \forall i, t$). Per-period costs $\bar{F}^1$ are set at 2.8% of permanent income and entry costs are set to zero.

Figure 1.67: Model without beliefs: hazard rate for exit under different per-period costs



Notes. This figure plots the hazard rate for exit in the model without beliefs ($b_{it} = 1\ \forall i, t$) for different levels of per-period participation costs $\bar{F}^1$. Entry costs are set to zero. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after entry to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified. The empirical hazard function is also shown in panel (B).

FIGURE 1.68: Model without beliefs: number of spells under different per-period costs

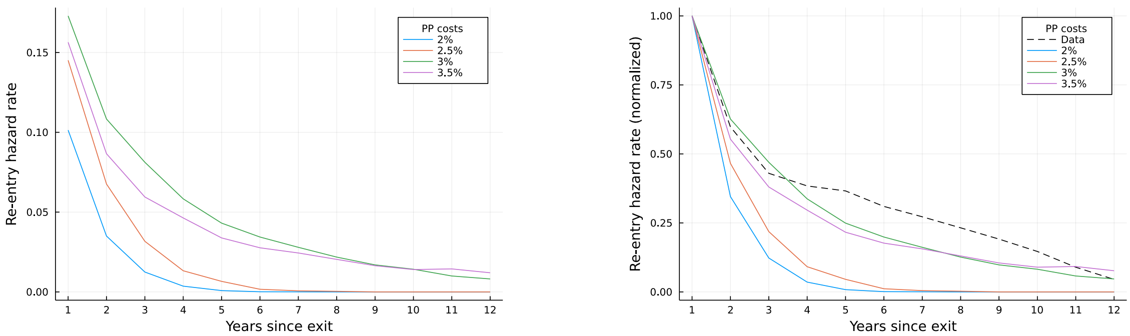


Notes. This figure plots the distribution of the number of spells in the simulated population for the model without beliefs ($b_{it} = 1 \; \forall i, t$) and under different levels of per-period participation costs. Entry costs are set to zero. The empirical distribution for the Norwegian population is also shown (see Figure 1.9).

FIGURE 1.69: Model without beliefs: hazard rate for reentry for different per-period costs

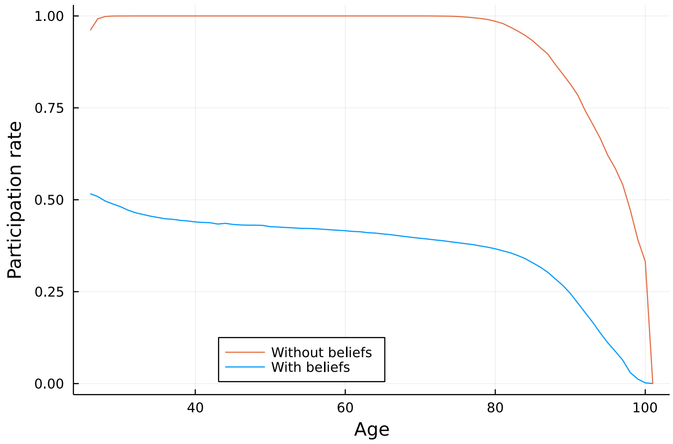
(A) Levels

(B) Slope (normalized)



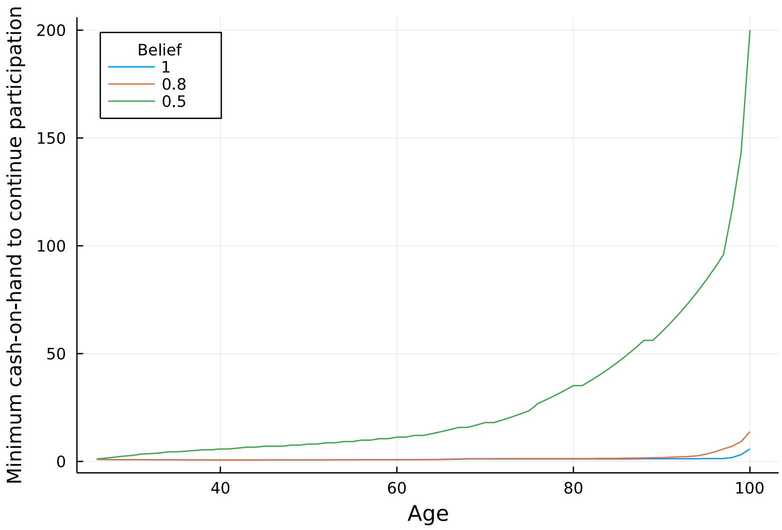
Notes. This figure plots the hazard rate for reentry in the model without beliefs ($b_{it} = 1 \; \forall i, t$) for different levels of per-period participation costs $\bar{F}^1$. Entry costs are set to zero. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after exit to 1 to facilitate comparison with the empirical hazard function, for which only the slope of the hazard function is identified. The empirical hazard function is also shown in panel (B).

Figure 1.70: Model with beliefs: simulated participation rates



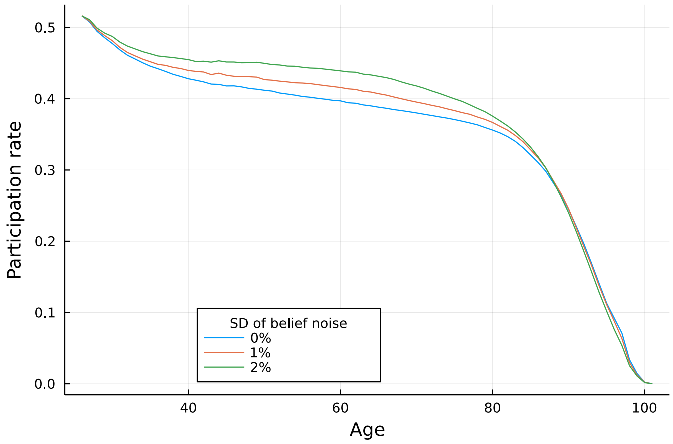
Notes. This figure plots the simulated participation rate over age in the models with and without beliefs. Entry and per-period costs are both set to 0.5% of permanent income in the two models.

Figure 1.71: Minimum wealth needed to continue participation for different beliefs



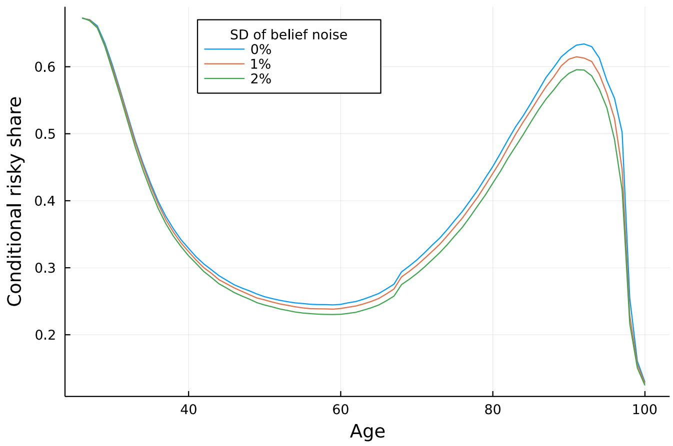
Notes. This figure plots the minimum wealth required to continue participating at different ages for three different values of beliefs $b_{it} \in \{0.5, 0.8, 1\}$.

FIGURE 1.72: Model with beliefs: simulated participation rates for different σ_ν



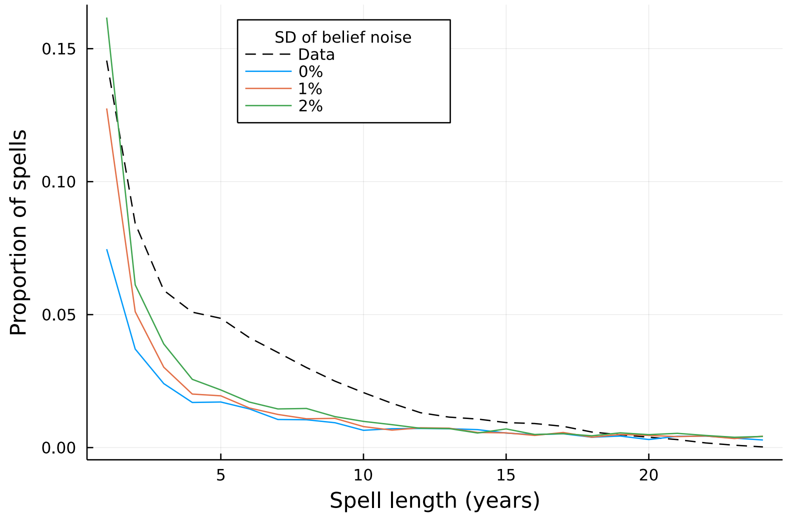
Notes. This figure plots the simulated participation rate over age in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income.

FIGURE 1.73: Model with beliefs: conditional risky asset share for different σ_ν



Notes. This figure plots the average conditional risky asset share over age in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income.

FIGURE 1.74: Model with beliefs: spell length distribution for different σ_ν

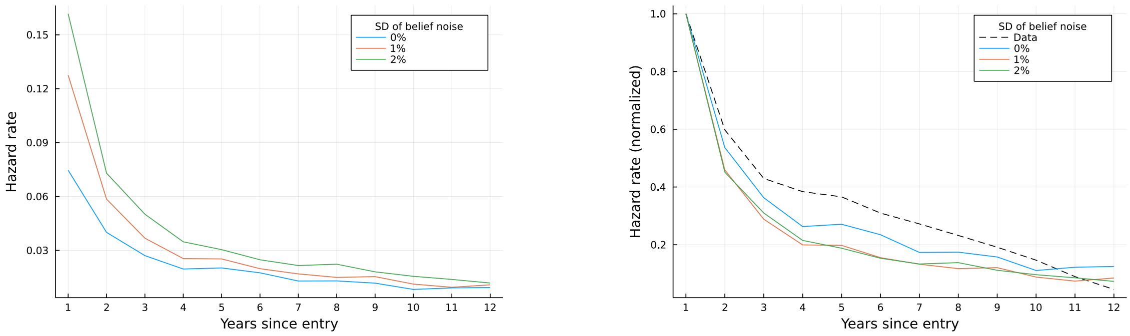


Notes. This figure plots the distribution of spell lengths in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income.

FIGURE 1.75: Model with beliefs: hazard rate for exit under different σ_ν

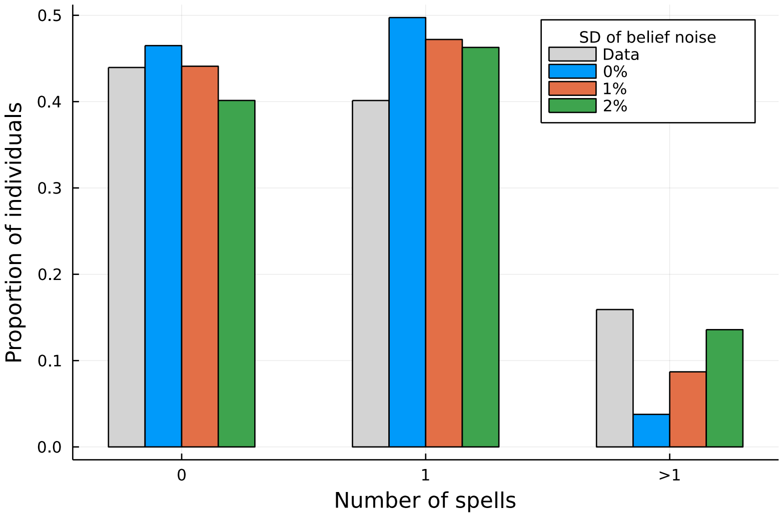
(A) Levels

(B) Slope (normalized)



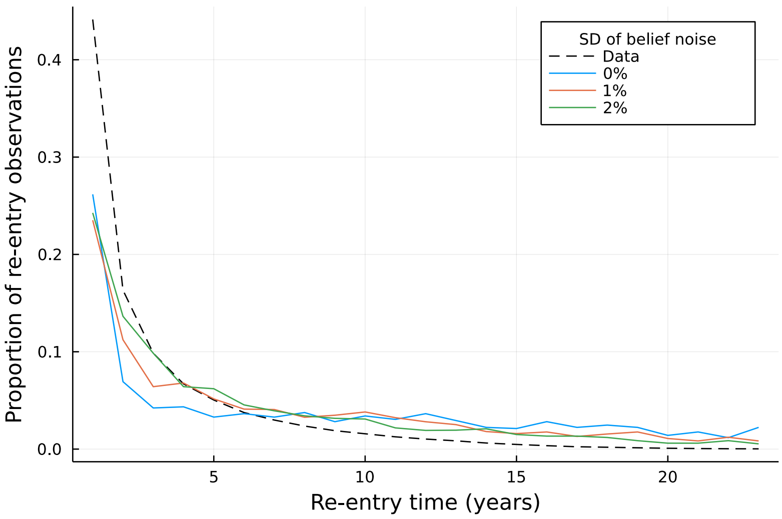
Notes. This figure plots the simulated hazard rate for exit in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after entry to 1, and shows the slope of the hazard function. The empirical hazard function is also shown in panel (B).

FIGURE 1.76: Model with beliefs: number of spells under different σ_ν



Notes. This figure plots the distribution of the number of spells in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income.

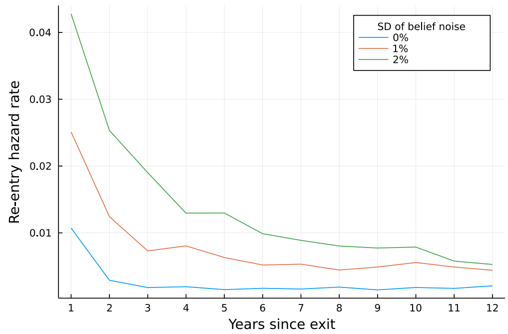
FIGURE 1.77: Model with beliefs: reentry time distribution under different σ_ν



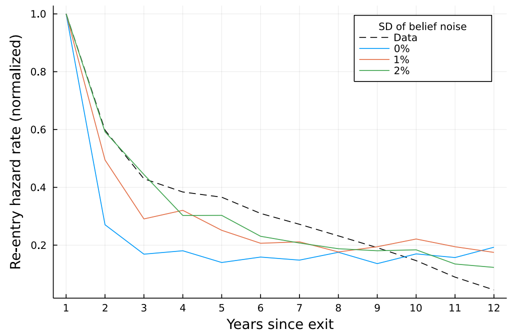
Notes. This figure plots the distribution of reentry times in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income.

Figure 1.78: Model with beliefs: hazard rate for reentry under different σ_ν

(A) Levels

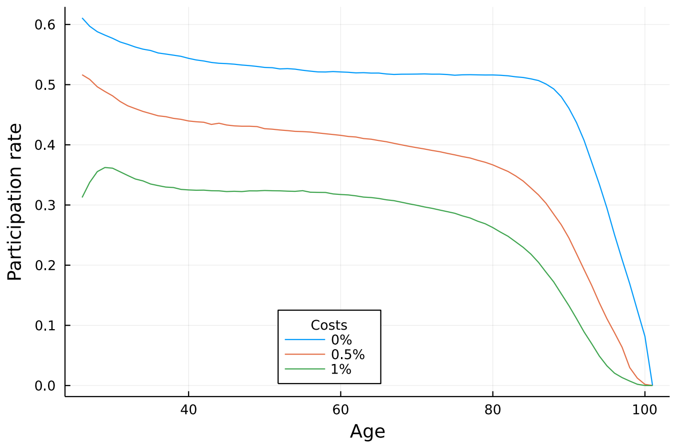


(B) Slope (normalized)



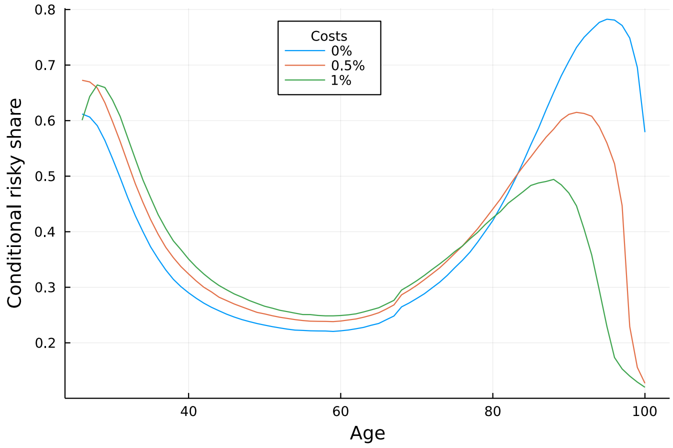
Notes. This figure plots the simulated hazard rate for reentry in the model with beliefs for different values of the standard deviation of belief shocks σ_ν. We consider 3 values: 0%, 1% (baseline), and 2%. Both entry and per-period costs are set at 0.5% of permanent income. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after exit to 1, and shows the slope of the hazard function. The empirical hazard function is also shown in panel (B).

Figure 1.79: Model with beliefs: simulated participation rates under different participation costs



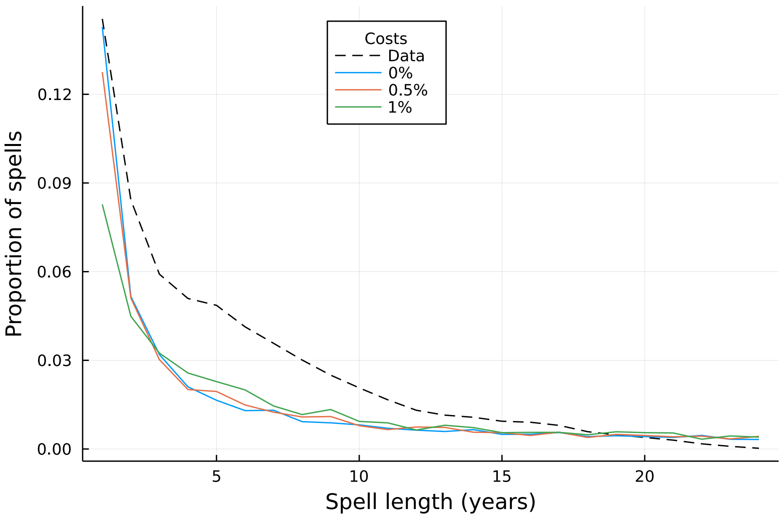
Notes. This figure plots the simulated participation rate over age in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs.

Figure 1.80: Model with beliefs: conditional risky asset share under different participation costs



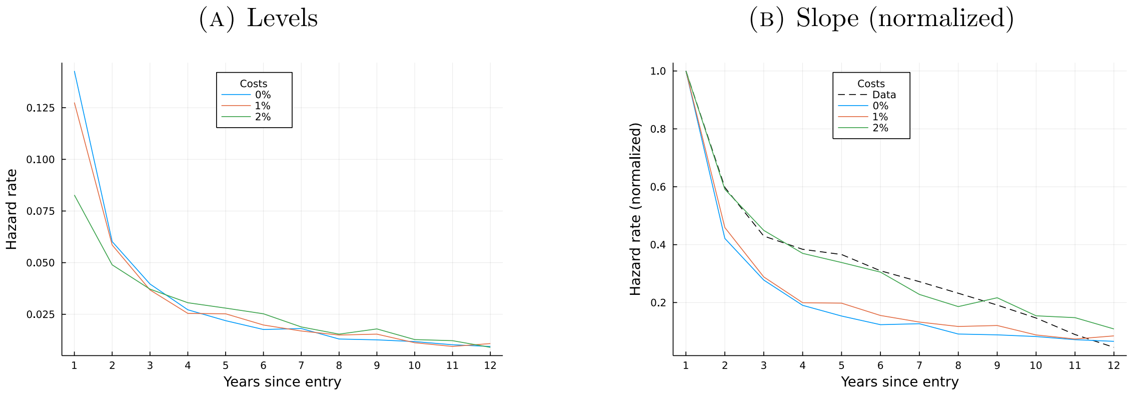
Notes. This figure plots the average conditional risky asset share over age in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs.

Figure 1.81: Model with beliefs: spell length distribution under different participation costs



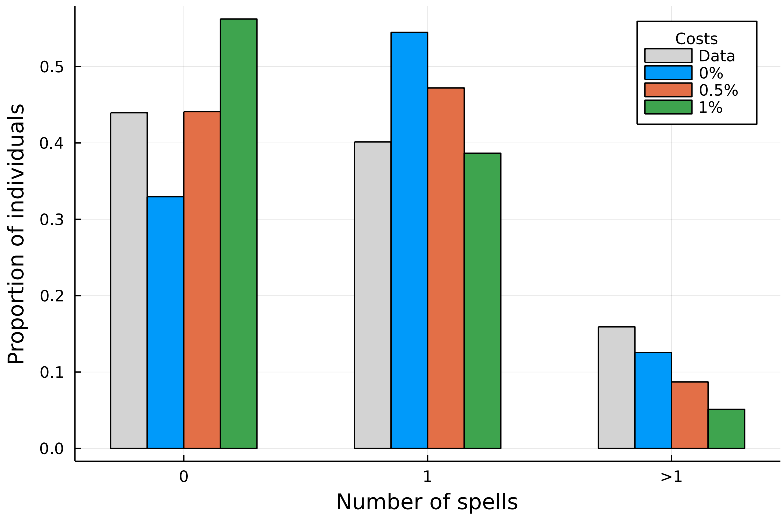
Notes. This figure plots the spell length distribution in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs.

Figure 1.82: Model with beliefs: hazard rate for exit under different participation costs



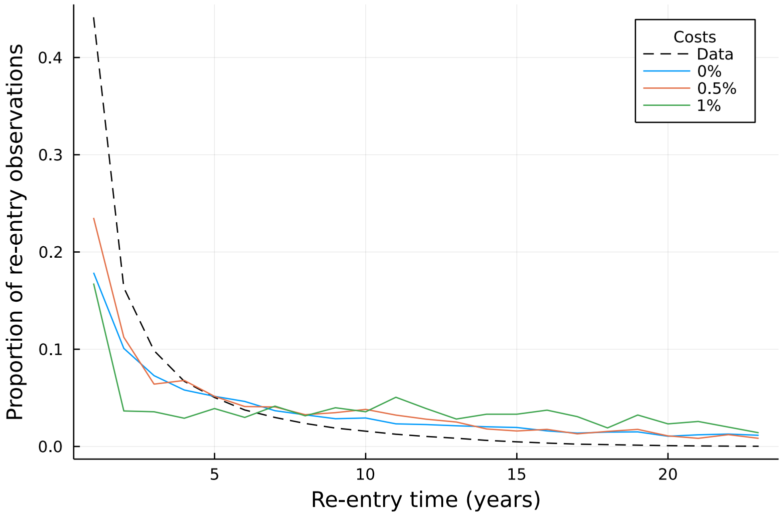
Notes. This figure plots the simulated hazard rate for exit in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after entry to 1, and shows the slope of the hazard function. The empirical hazard function is also shown in panel (B).

Figure 1.83: Model with beliefs: number of spells under different participation costs



Notes. This figure plots the distribution of the number of spells in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs.

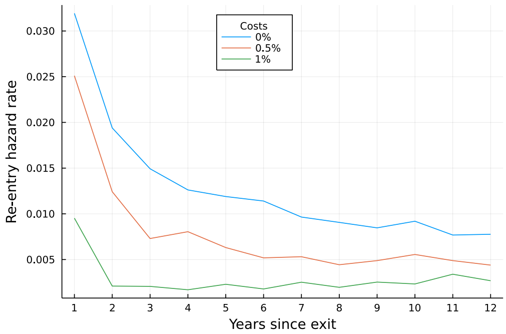
FIGURE 1.84: Model with beliefs: reentry time distribution under different participation costs



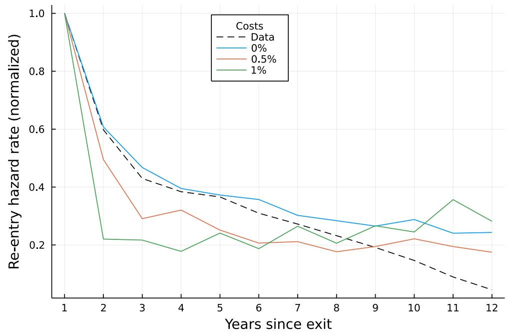
Notes. This figure plots the distribution of reentry times in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs.

FIGURE 1.85: Model with beliefs: hazard rate for reentry under different participation costs

(A) Levels

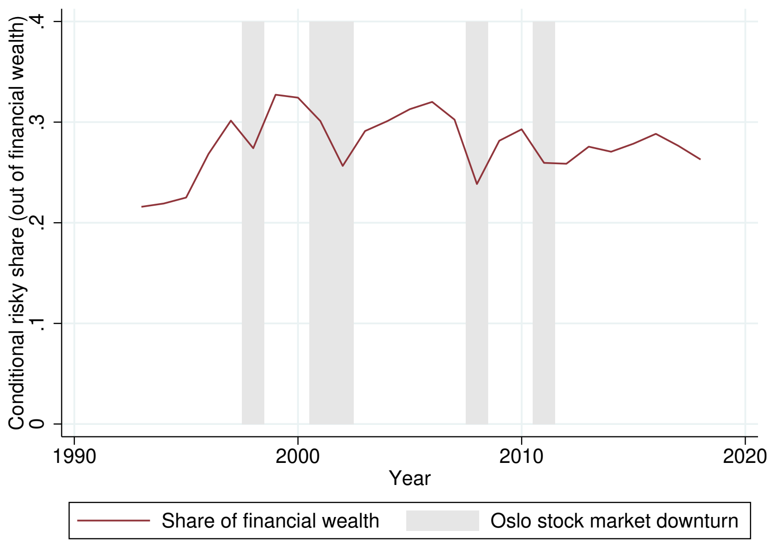


(B) Slope (normalized)



Notes. This figure plots the simulated hazard rate for reentry in the model with beliefs under different values of participation costs. We consider 3 values: 0%, 0.5% (baseline), and 1%. We apply this value to both entry and per-period costs. Panel (A) gives the hazard rates in levels in each case. Panel (B) normalizes the hazard rate in the year after exit to 1, and shows the slope of the hazard function. The empirical hazard function is also shown in panel (B).

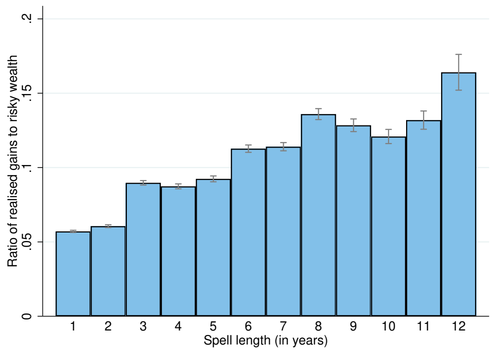
Figure 1.86: Average conditional risky share over time



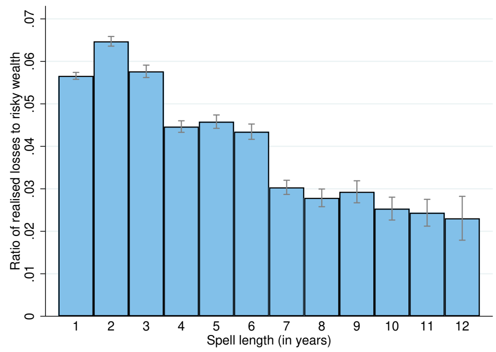
Notes. This figure plots the average conditional risky share over time. The shaded areas are stock market downturn years in which the Oslo Børs Benchmark Index fell by at least 10%.

Figure 1.87: Performance of exiters by spell length

(A) Gains scaled by risky wealth

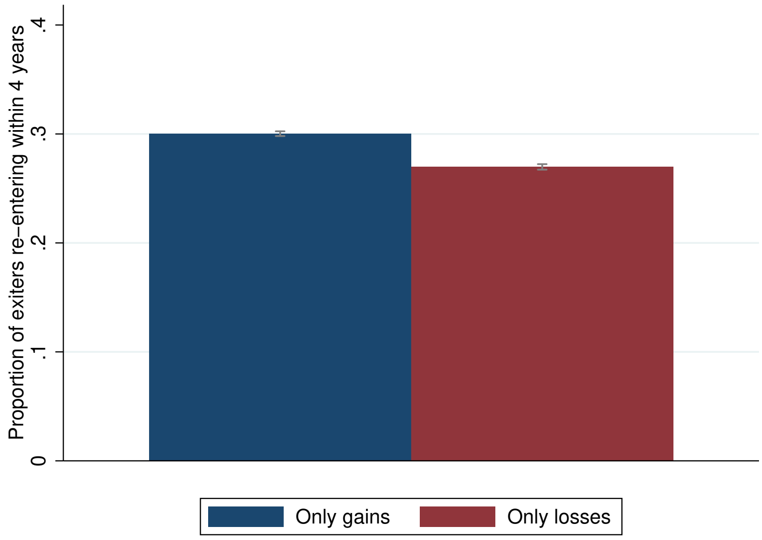


(B) Losses scaled by risky wealth



Notes. This figure shows the performance of exiters by spell length based on records of taxable gains and tax-deductible losses in the income tax data. In panel (A), we plot the average ratio of taxable gains in the year of exit to risky financial wealth in the year before exit for exiters of different spell lengths. Taxable gains from the sale of stocks and funds is computed as the sum of items TR 3.1.8, TR 3.1.9, and TR 3.1.10 in the tax records. In panel (B), we plot the equivalent ratio for losses, which are computed as the sum of items TR 3.3.8, TR 3.3.9, and TR 3.3.10. We use exiters who enter from 2006 onward in these plots.

Figure 1.88: Proportion of exiters reentering within 4 years by prior performance



Notes. This figure plots the proportion of exiters who reenter into the stock market within the next 4 years based on their prior performance as measured by the report of taxable gains and tax-deductible losses. The left bar shows the proportion of exiters who only report taxable gains in their exit year who reenter in the next 4 years. The right bar shows the corresponding proportion for exiters reporting only losses. We use exiters who exit between 2006 and 2014 in these plots.

Chapter 2

The impact of changes in bank capital requirements

1 Introduction

The Global Financial Crisis of 2008 highlighted multiple vulnerabilities within the banking sector (Basel Committee on Banking Supervision, 2009; Sironi, 2018). In order to mitigate these financial stability risks, regulators have since employed various micro- and macro-prudential policy tools, one of which is bank capital requirements. These requirements, typically set as a minimum ratio of total regulatory capital to risk-weighted assets, aim to ensure that banks can withstand unexpected losses and maintain solvency in a crisis. Banks can respond to capital regulation in various ways, and the choice of response could have different macroeconomic and financial stability implications (Hanson et al., 2011). With respect to an increase in requirements, banks can accumulate more capital, reduce total assets or shift their asset composition towards less risky assets. They could also simply maintain their capital ratios and dig into their pre-existing capital buffer provided this buffer is sufficiently large. If banks lower lending as part of their adjustment, this could adversely affect macroeconomic activity today. Instead, accumulation of more capital can improve bank resilience to future shocks, thus improving financial stability. This paper seeks to shed light on the adjustment of banks to capital regulation using confidential regulatory returns data for UK banks.

Estimating the impact of capital requirements poses empirical challenges: first, in most countries, bank capital requirements are homogeneously set across banks, often at the Basel minimum of an 8% risk-based capital ratio, which leaves little, if any, cross-sectional variation to exploit for identification. When cross-sectional variation is available, studies are sometimes constrained to look at one-off regulatory changes and to compare "treated" and "untreated" banks around singular events (e.g., Mésonnier and Monks, 2015; Gropp et al., 2018). Focusing on isolated regulatory changes can constrain the time dimension and make it difficult to study effects at longer horizons. If the policy change is particularly unique or targeted at specific banks, it may be difficult to apply these results in other settings. The second challenge is that capital requirements are not randomly allocated, making it difficult to separate the effect of a change in capital requirements from the fact that banks receiving a higher or lower requirement may be inherently different from those that do not. This selection problem can lead to endogeneity concerns if the regulatory change is not orthogonal to other drivers of the outcome.

In this paper, I address each of these empirical concerns: first, the UK is a unique setting because time- and bank-varying capital requirements known as *trigger ratios* have been in place for all regulated banks since 1989, thus providing a long time dimension simultaneously with cross-sectional variation. On the second empirical challenge, existing papers that study UK requirements reference anecdotal evidence to argue that regulators focus on non-balance sheet risks such as organizational structures and reporting procedures when setting trigger ratios (Aiyar et al., 2014). I apply the least absolute shrinkage and selection operator (lasso) of Tibshirani (1996) to provide statistical evidence for this argument. This finding supports the identification assumption applied in the analysis that changes in bank-level capital requirements can be treated as orthogonal to bank balance sheet risks.

Using local projections à la Jordá (2005), I find that bank capital ratios do respond to a change in required ratios, although the pass-through is less than one-for-one. In my baseline specification, a 1 percentage point (pp) increase in trigger ratios causes a 0.5pp rise in actual capital ratios. Adjustments of capital ratios occur primarily through two channels: the first is capital accumulation, whereby total regulatory capital increases by around 1% in the year following a 1pp increase in capital requirements. This is predominantly driven by Tier 2 capital, which rises by 3-4% during this period. The second is a risk composition effect, whereby banks adjust their asset portfolios towards less risky assets with average risk weights, computed as the ratio of risk-weighted assets and total assets, falling by 1-1.5pps during the three years following the regulatory change. There is no significant effect on bank lending. By splitting the sample based on the direction of the regulatory change, I show that bank capital ratios respond to decreases, but not increases, in capital requirements. Instead, banks opt to dig into their existing capital buffers when faced with tighter requirements. The response of banks also appears to have changed since the financial crisis. I find that the risk composition effect is a post-crisis result. Prior to the financial crisis, adjustments occurred mainly through capital accumulation, in particular Tier 2 capital, as well as lending with the quantity of loans dropping by 5% one year after a 1pp rise in requirements.

The remainder of the paper is organized as follows: Section 2 gives an overview of the existing literature and Section 3 gives an account of the UK regulatory framework. Section 4 describes the data used in the empirical analysis with Section 5 providing descriptive statistics. Section 6 outlines the methodology and Section 7 shows the results. Robustness checks are provided in Section 8, while Section 9 concludes.

2 Literature review

The theoretical discussion gives three conditions that must be satisfied for changes in capital requirements to impact bank lending and balance sheet composition.[1] The first condition is that capital requirements should be effectively binding. This does not necessarily mean that capital ratios must exactly equal the required level at all times. Banks may instead have a desired capital buffer in excess of their requirement that they wish to keep constant. The second condition is that credit demand cannot be inelastic to allow for loan quantities to adjust following a regulatory change. Third, acquiring additional capital should be more expensive relative to debt. For this to be the case, the Modigliani-Miller theorem must not hold. The theorem states that if there are no frictions, changes in the composition of a bank's liabilities have no effect on funding costs, and as a result should have no effect on bank lending (Modigliani and Miller, 1958).[2] Kashyap et al. (2010) calibrate a model based on the Modigliani and Miller (1958) framework, in which the main difference in the cost of equity and debt financing is differential tax treatments. They find modest long-run impacts of higher capital requirements on lending rates with the cost of borrowing rising by only 25-45 basis points following a 10pp increase in capital requirements. Elliott (2009) also finds small impacts of capital requirements on loan volumes of US banks, while Miles et al. (2013) reach the same conclusion for UK banks.

Within the empirical literature, one category of papers has studied the response of banks

[1]For an overview of the theoretical literature on the impact of capital requirements, see VanHoose (2007, 2008).

[2]Examples of frictions are tax deductibility of debt (Modigliani and Miller, 1958) and asymmetric information that makes it costly to raise external equity (Myers and Majluf, 1984). Equity capital may also be more expensive due to ex-post verification costs (Diamond, 1984; Gale and Hellwig, 1985).

to a capital requirement change.[3] My paper falls into this strand of the literature. Using a fixed effects framework, Aiyar et al. (2014) show that a 1pp rise in bank capital requirements is associated with a 5.7-8% decline in bank lending in the subsequent three quarters. Ediz et al. (1998) find, using a dynamic multivariate panel regression model and data from 1989-1995, that bank capital ratios do react to changes in required ratios, though much of the reaction is through adjusting capital rather than loan quantities. Bridges et al. (2014) study the impact of capital requirements on bank lending using UK data from 1989-2011. Using dynamic panel regressions, they find that changes in capital requirements do affect capital ratios. Following a tightening of requirements, banks rebuild their buffers by increasing their capital ratios over time. The authors also find heterogeneous responses of bank lending across sectors with commercial real estate lending growth showing the largest decline, followed by other corporate lending and then household secured lending. My paper builds on this by studying whether banks respond in other ways, in particular via capital accumulation or the risk composition of the asset portfolio. I use local projections à la Jordá (2005) to allow for greater flexibility in the shape of the impulse responses, and have a longer time dimension that allows for comparison of pre- and post-financial crisis responses. I also statistically test, using lasso methods, an assumption implicitly made in Bridges et al. (2014) and motivated by anecdotal evidence given in Aiyar et al. (2014) that changes in requirements can be treated as exogenous with respect to balance sheet risks and thus are orthogonal to other drivers of bank lending.

Francis and Osborne (2012) follow a different empirical approach initially introduced by Hancock and Wilcox (1993, 1994), and estimate a partial adjustment model whereby banks have a target capital ratio that depends on the regulatory requirement amongst other factors. Due to adjustment costs, they cannot adjust instantly or fully to their new target ratio. Using UK data, the authors find small effects of capital requirements on lending. de Ramon et al. (2022) apply this method to compare the pre- and post-crisis responses of banks to capital requirements. In line with our findings, they show that before the crisis, banks responded to changes in requirements via reductions in loan quantities and accumulation of capital, in particular Tier 2 capital. They show that banks have focused on capital accumulation as their primary adjustment tool since the financial crisis. We find instead that banks have shifted to adjusting the risk composition of their assets.[4]

[3]There is also a literature looking at the impact of capital shocks not driven by regulation (e.g., Bernanke et al., 1991; Peek and Rosengren, 1997; Heid et al., 2004; Fonseca and González, 2010; Jiménez et al., 2010; Stolz and Wedow, 2011).

[4]Papers that have studied capital requirements outside of the UK include Mésonnier and Monks (2015), Jiménez et al. (2017), Fang et al. (2018), Gropp et al. (2018) and De Jonghe et al. (2019).

3 Institutional background

An appealing feature of the UK regulatory regime is that since 1989, supervisors have set bank- and time-varying minimum capital requirements in excess of the 8% requirement given by the Basel Accords.[5] The variation in the magnitudes and timing of capital requirement changes across banks, in addition to the fact that discretionary policy has been a feature of the UK supervisory regime for many years, makes the UK an appealing setting for studying the impact of capital requirements.

From 1997-2001, supervisors followed the Risk Assessment, Tools and Evaluation (RATE) framework (Financial Services Authority, 1998). It had three key stages as shown in Figure 2.8: an initial formal risk assessment, a risk mitigation supervisory programme and the evaluation of the supervisory actions and outcomes. The risk assessment was based on nine evaluation factors that can be grouped into one of two categories: business risk and control risk. Business risk covered six quantitative factors and involved an analysis of the bank's financial position and key business.[6] Control risk determines the adequacy of the internal control framework and covers the remaining three qualitative factors.[7] Following an assessment of business and control risks, a supervisory programme was sent to the bank outlining the regulator's concerns and providing a set of actions that could include a new capital requirement. As such, a wide range of risks, both balance sheet and non-balance sheet risks, were covered within the RATE framework.[8] The resulting capital requirement, set as a proportion of risk-weighted assets, was known as the *trigger ratio*.

In 2001, the FSA replaced RATE with ARROW (Advanced Risk Responsive Operation frameWork). An important difference of the ARROW framework relative to RATE is that under ARROW, the FSA followed a Risk to Our Objectives (RTO) approach, whereby the risk of interest to the FSA was not commercial risk taking per se, but rather the risk that the FSA's four statutory objectives would not be met. Indeed, "*it is not the role of the FSA to restrict appropriate risk-taking by regulated institutions or investors*" (Financial Services Authority, 2000, p. 4). The four objectives were: maintaining confidence in the

[5] Basel I introduced minimum capital requirements, whereby banks were required to satisfy a ratio of total regulatory capital to total risk-weighted assets of 8%, half of which needed to come from Tier 1 capital (Basel Committee on Banking Supervision, 1988). Iterations of the Basel Accords have since brought in changes to capital regulation. For details on Basel II and III, see Basel Committee on Banking Supervision (2006) and Basel Committee on Banking Supervision (2010a,b) respectively.

[6] The six quantitative factors are capital, asset quality, market risk, earnings, liabilities and liquidity profile, and business risk profiles.

[7] The three qualitative factors are internal controls, organizational structure and management.

[8] The intensity of the supervisory relationship was higher, the greater the perceived risk profile of a bank. The length of time between formal risk assessments was smaller at approximately 6-12 months for banks with high perceived risks compared to 18-24 months for banks with low risk profiles (Financial Services Authority, 1998). Figure 2.9 illustrates this concentration of resources towards "riskier" banks.

UK financial system, promoting public understanding of the financial system, securing the appropriate degree of protection for consumers and reducing the scope for financial crime. As with RATE, business and control risks were evaluated and used for risk mitigation programmes that could include changes in capital requirements.[9] The Prudential Regulation Authority (PRA) was given responsibility over supervision in 2013.[10]

From this, it is clear that through the inclusion of control risks and the RTO approach of the FSA that capital requirement decisions were not based purely on balance sheet risks. There has been some anecdotal evidence discussed in Aiyar et al. (2014) suggesting that capital requirement decisions were mainly based on control risks, particularly in the pre-crisis era. The Turner Review stated that "*risk mitigation programs set out after ARROW reviews tended to focus more on organisation structures, systems and reporting procedures, than on overall risks in business models*" (Financial Services Authority, 2009, p. 87). Furthermore, "*under ARROW I there was no requirement on supervisory teams to include any developed financial analysis in the material provided to ARROW Panels*" (Financial Services Authority, 2008, p. 3). From this anecdotal evidence, it appears that capital requirement changes were orthogonal to balance sheet risks. I later provide statistical evidence using lasso regressions to support this. This institutional feature gives support to the identification assumption used in this paper that changes in capital requirements can be treated as exogenous with respect to other drivers of bank lending and balance sheet composition.

4 Data

This paper uses the *Historical Banking Regulatory Database* (HBRD) constructed in de-Ramon et al. (2017). By extracting information contained in mandatory regulatory returns, HBRD contains balance sheet and confidential regulatory information for all regulated banks and building societies in the UK. The data is provided at both a consolidated/banking group level and a solo bank level, and spans 1989H1 to 2013H2.[11] Following Bridges et al. (2014) and de Ramon et al. (2022), I use the consolidated dataset for the analysis as lending and capital decisions are typically made at the banking group level.

[9] The FSA devoted attention and resources to the high impact banks (typically larger banks) as they were perceived to pose the greatest potential threat to the FSA's objectives (Financial Services Authority, 2002; International Monetary Fund, 2003). Consequently, the probability assessment was not undertaken for low impact banks and they did not receive a risk mitigation programme.

[10] For further details on the PRA's supervisory framework, see Bank of England (2018).

[11] For building societies, capital requirements data begins in 1997.

The use of HBRD is especially beneficial for this work for a number of reasons: first, the dataset contains confidential information on individual bank capital requirements for the entire UK banking system. Second, the long time dimension allows for analysis of the medium-term impacts of capital requirements rather than focusing on the immediate-term response of banks. It also enables me to study whether bank responses to capital requirements have changed since the financial crisis. Third, HBRD contains over 100 analytical measures constructed using over 500 regulatory report items. The wide range of variables provides a large amount of information about each bank that would have been observable to the regulator and, as noted in Section 3, could be used when assessing bank risks and deciding on capital requirements. As such, this data is useful to statistically test whether balance sheet variables affect the regulator's capital requirement decision.[12]

A concern with the original raw dataset is the unbalanced nature of the panel associated with missing values and the entry/exit of banks throughout the sample. The raw dataset has 4,616 observations. I clean the original dataset using the steps described in Section B.1. The final dataset consists of 3,256 observations. Tables 2.3 and 2.4 provide details on the construction of key ratios and quantities used in the analysis.

5 Descriptive statistics

Table 2.1 provides summary statistics from the full sample. The average trigger ratio is 11.6%, which illustrates the use of discretionary capital requirements above the Basel I minimum of 8% by UK regulators. The high standard deviation of trigger ratios indicates the large cross-sectional variation in trigger ratios across banks. This is further highlighted in Figure 2.1, which shows the distribution of trigger ratios over time. The largest trigger ratios declined in the years building up to the Great Recession and then increased in the years following it.

In terms of capital requirement changes, Table 2.1 shows that there are 606 occurrences of capital requirement changes in the sample, making this almost a one-in-five event.[13] Although the median change is negative, the mean is slightly positive, suggesting that increases in capital requirements tend to be larger in magnitude than decreases. This is reinforced in Figure 2.2, which shows the distribution of capital requirement changes.

[12] One concern when using regulatory returns data over such a long period is changes in reporting frameworks and variable definitions. When constructing the HBRD, de-Ramon et al. (2017) use the instructions from each framework to construct consistent measures of variables over time.

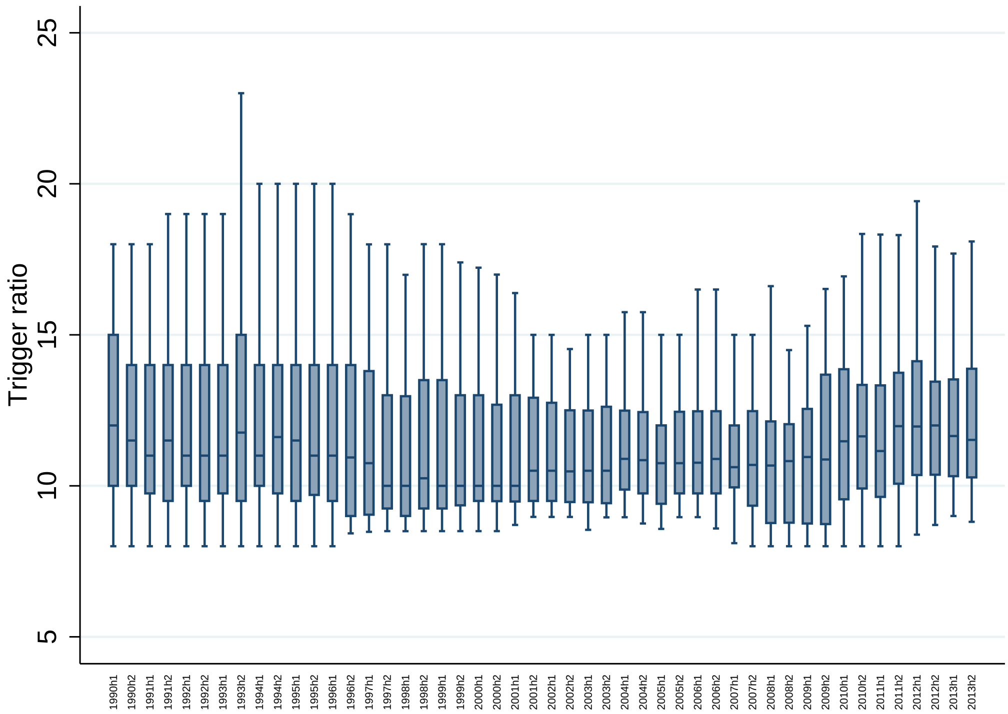
[13] I classify a change as having occurred if the absolute value of the half-year change in trigger ratios exceeds 0.1pps.

TABLE 2.1: Summary statistics

Variable	Obs	Mean	Std. Dev.	P10	P50	P90
Trigger ratio	3256	11.587	2.767	8.958	10.964	15
Change in trigger ratio (half-year, if change)	606	.045	1.181	-1.107	-.109	1.36
Tier 1 risk-based capital ratio	3256	19.22	16.078	7.353	12.896	43.414
Total risk-based capital ratio	3256	21.511	14.613	10.918	16.199	41.948
Capital buffer	3256	9.831	13.347	1.331	4.473	28.318
Assets growth (half-year)	3256	3.218	9.199	-8.584	3.57	14.553
Risk-weighted assets growth (half-year)	3256	2.807	9.005	-8.956	3.356	13.43
Average risk weight	3256	56.902	20.382	29.807	53.684	87.971
Liquid asset ratio	2829	10.731	10.722	.163	8.142	26.255
Tier 1 leverage ratio	3256	10.551	9.35	3.819	6.666	25.441
Solvency ratio	3256	179.306	98.471	113.293	141.812	317.349
Loans-to-assets ratio	3256	50.877	27.618	10.472	53.132	84.543
Tier 1 capital growth (half-year)	3255	3.193	8.334	-6.071	2.358	13.555
Total capital growth (half-year)	3256	3.047	8.038	-6.292	2.494	13.586
Loans growth (half-year)	3256	3.095	11.133	-11.24	3.655	16.735
Deposits growth (half-year)	3227	2.952	16.534	-13.88	3.196	20.294
Unsecured loans growth (half-year)	3155	4.583	33.645	-15.72	2.668	20.998
Residential loans growth (half-year)	2621	11.593	200.936	-17.783	3.367	25.698

Note: this table provides summary statistics based on the full sample. Columns 1-3 give the number of observations, mean and standard deviation for each variable. Columns 4-6 show the 10^{th}, 50^{th} and 90^{th} percentiles. “Trigger ratio” is the proportion of risk-weighted assets that banks must hold as capital. “Change in trigger ratio (half year, if change)” is the half-year change in the trigger ratio using only those observations where a capital requirement change occurred (a change is coded as having occurred if the half-year change exceeds 0.1pps in absolute value). “Average risk weight” is the ratio between risk-weighted assets and total assets. “Liquid asset ratio” is the ratio of liquid assets to total assets, where liquid assets here are defined as high quality liquid assets as well as credit to other financial institutions, debt securities and equity shares. “Tier 1 leverage ratio” is the ratio of Tier 1 capital to total assets. “Solvency ratio” is the ratio between total regulatory capital and total required capital, where total required capital is given by the trigger ratio multiplied by risk-weighted assets. “Unsecured loans growth” is the half-year growth of loans not secured on residential property.

Figure 2.1: Box plot of trigger ratios over time

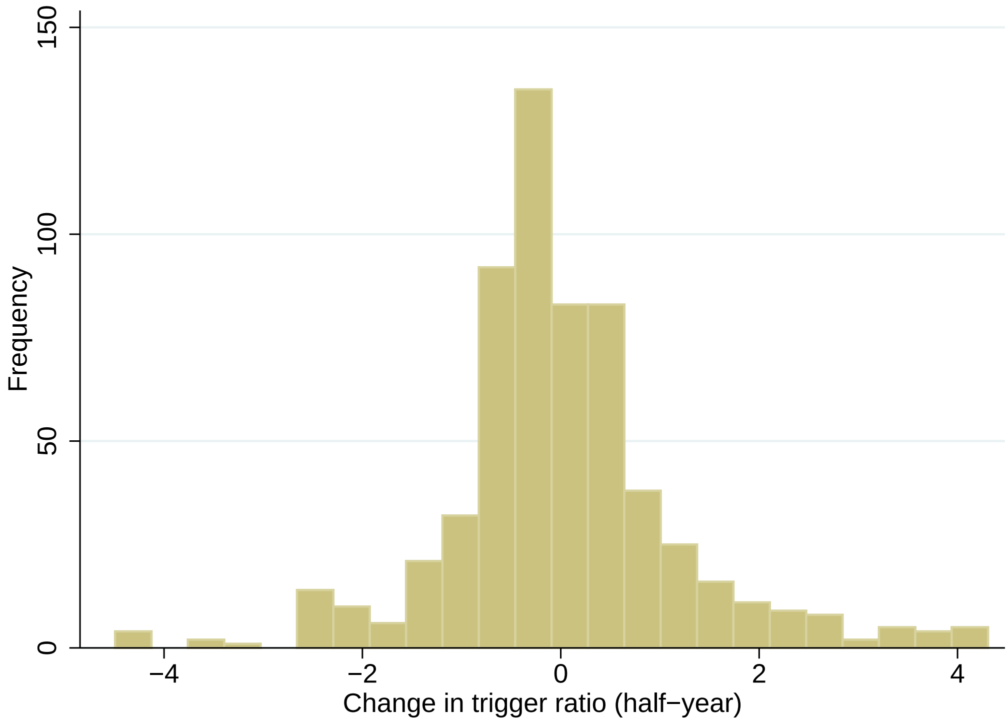


Note: this figure shows the distribution of trigger ratios over time. The points correspond to the lower adjacent value, 25$^{\text{th}}$ percentile, median, 75$^{\text{th}}$ percentile and upper adjacent value.

From this histogram, the rightward skew is clear. In order to see whether capital requirement changes tend to occur simultaneously across banks, Figure 2.3 plots the proportion of banks experiencing a change in their trigger ratio over time. While there are some periods when no trigger ratio changes occurred, namely pre-1995H2 and 2005H1-2006H2, there are regulatory changes in every other period, thus suggesting that capital regulation was active throughout the sample rather than only in specific periods. The frequency of trigger ratio movements appears to have increased since the Great Recession with around 50% of banks experiencing a change in each half year during the post-crisis period compared to less than 20% in most pre-crisis periods. A concern may be that trigger ratios move in the same direction for all banks experiencing a change. This could suggest that regulators are responding to business cycle fluctuations rather than individual bank characteristics. As noted in Meeks (2017) and shown in Figure 2.4, there are few periods where changes in capital requirements are of the same sign for all banks experiencing a regulatory change. Figure 2.4 also shows that the spread of trigger ratio changes has risen since the crisis. The fact that the post-crisis period does not show purely positive changes in capital requirements indicates that increased supervisory attention rather than just tighter microprudential policy is a feature of the post-crisis period.

Although most UK banks face trigger ratios in excess of the 8% Basel I minimum risk-based

Figure 2.2: Distribution of capital requirement changes



Note: this histogram shows the frequency of half-year changes in trigger ratios across a number of narrow bins for the full sample 1989H1-2013H2. I exclude observations with absolute changes of less than 0.1pps.

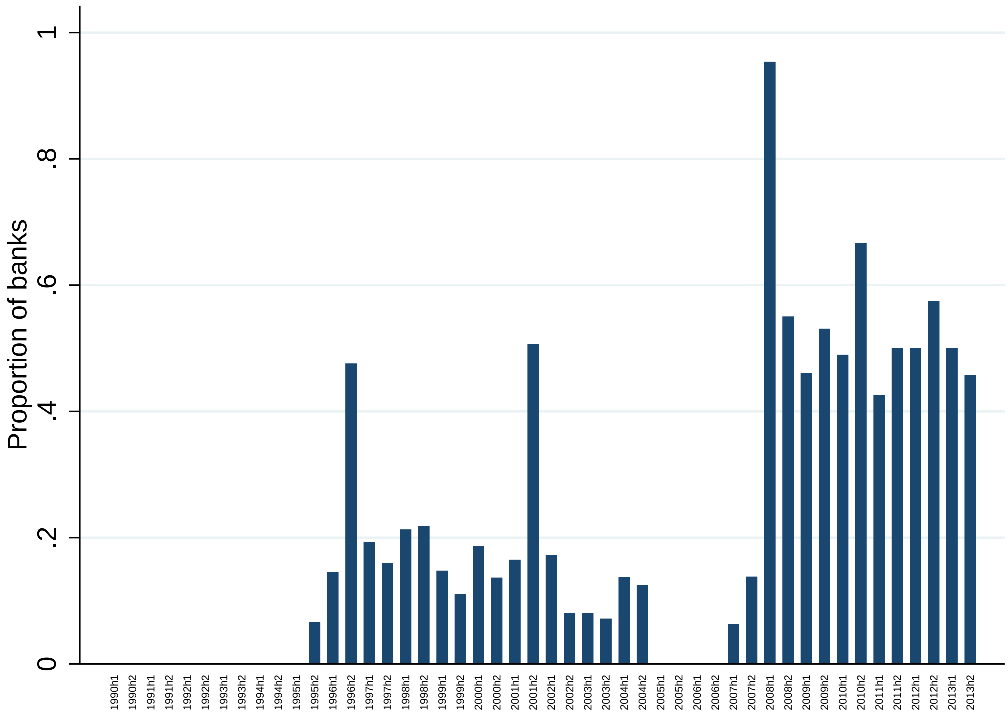
capital ratio, many still hold capital buffers in excess of their requirements. Table 2.1 shows that the average capital buffer - computed as the difference between the risk-based capital ratio and the trigger ratio - is 9.8pps.[14] Figure 2.5 plots the distribution of capital buffers over time. Barring some banks in the first few periods of the sample, banks did not fall short of their required capital ratios, suggesting that capital regulation has been enforced. There is a large dispersion in capital buffers across banks with some banks operating close to their requirements and others holding substantial buffers.[15] It appears rare that a bank would operate with a capital ratio exactly equal to their requirement. A feature of this box plot is that the distribution of capital buffers became much more concentrated in the years leading up to the Great Recession as banks originally holding the largest buffers reduced them. This could indicate countercyclicality of capital buffers.[16] Some evidence for countercyclicality of buffers is provided in Figure 2.6, which plots aggregate banking sector trigger and capital ratios over time. There was a slight decline in buffers from

[14] Other papers have also shown that banks tend to have capital ratios in excess of regulatory minima, both in the UK and in other countries (e.g., Lindquist, 2004; Jokipii and Milne, 2008; Shim, 2013).

[15] One reason for holding capital buffers is to avoid breaching capital requirements (see Alfon et al., 2004, Peura and Keppo, 2006 and Francis and Osborne, 2012).

[16] Some evidence for countercyclicality is provided in Stolz and Wedow (2011) for German banks, Ayuso et al. (2004) for Spanish banks and Shim (2013) for US banks. However, there are also papers giving evidence of procyclicality of capital buffers, e.g., Montagnoli et al. (2018) for Portuguese banks and Valencia and Bolaños (2018) for developing countries.

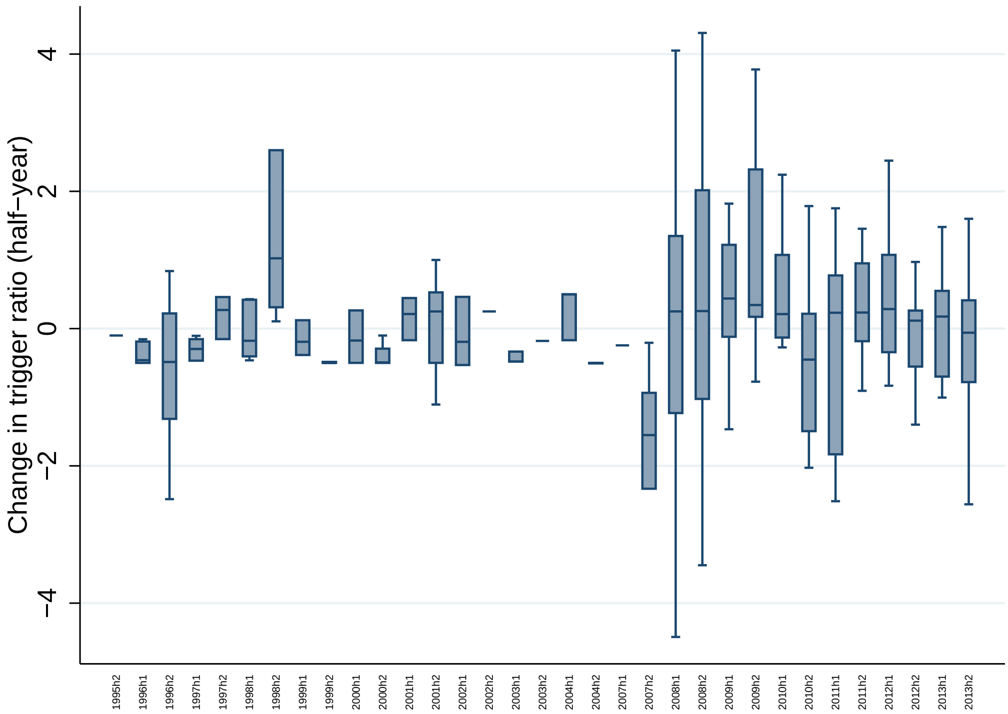
Figure 2.3: Proportion of banks experiencing changes in trigger ratios over time



Note: this figure shows the proportion over time of banking groups who experienced a change in trigger ratios in a given period. A change is coded as having occurred if the absolute change in trigger ratios is more than 0.1pps relative to the previous half year.

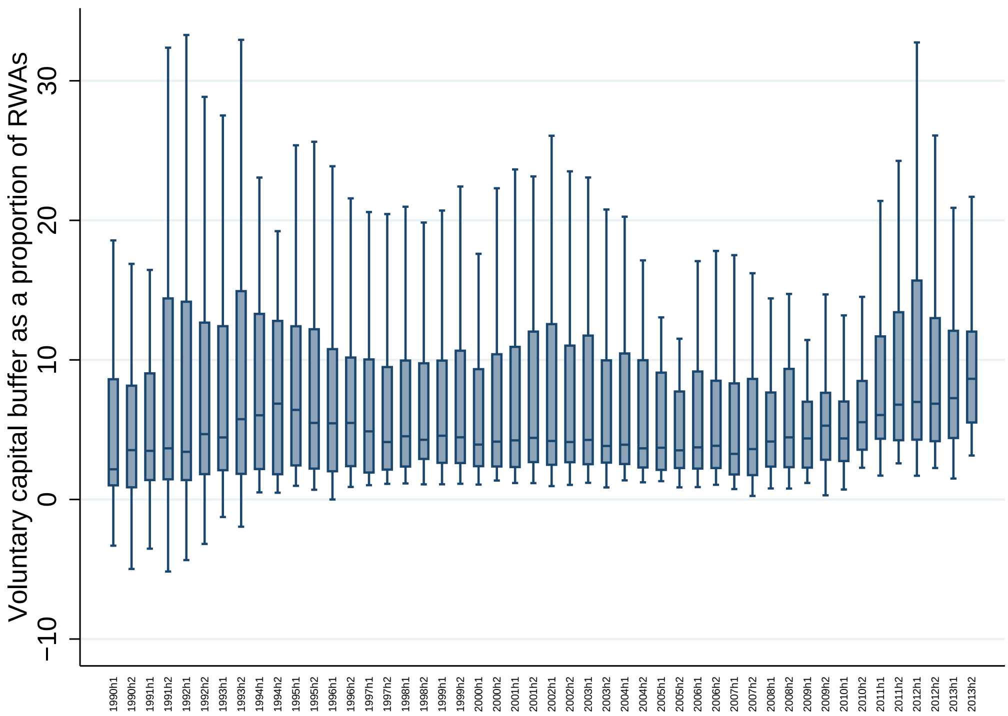
the 2000s and then a sharp rise during the recovery phase following the financial crisis. However, it should be noted that the rise in buffers coincides with increasing supervisory attention following the Great Recession and so perhaps cannot be completely attributed to business cycle fluctuations. A notable takeaway from the figure is that aggregate trigger ratios have been very stable at just under 10% until around 2010, after which there was a small rise to just over 10%, a feature also documented in de-Ramon et al. (2017). This suggests that much of the action of capital requirements in the UK has been at the micro rather than macro level.

Figure 2.4: Box plot of changes in trigger ratios over time



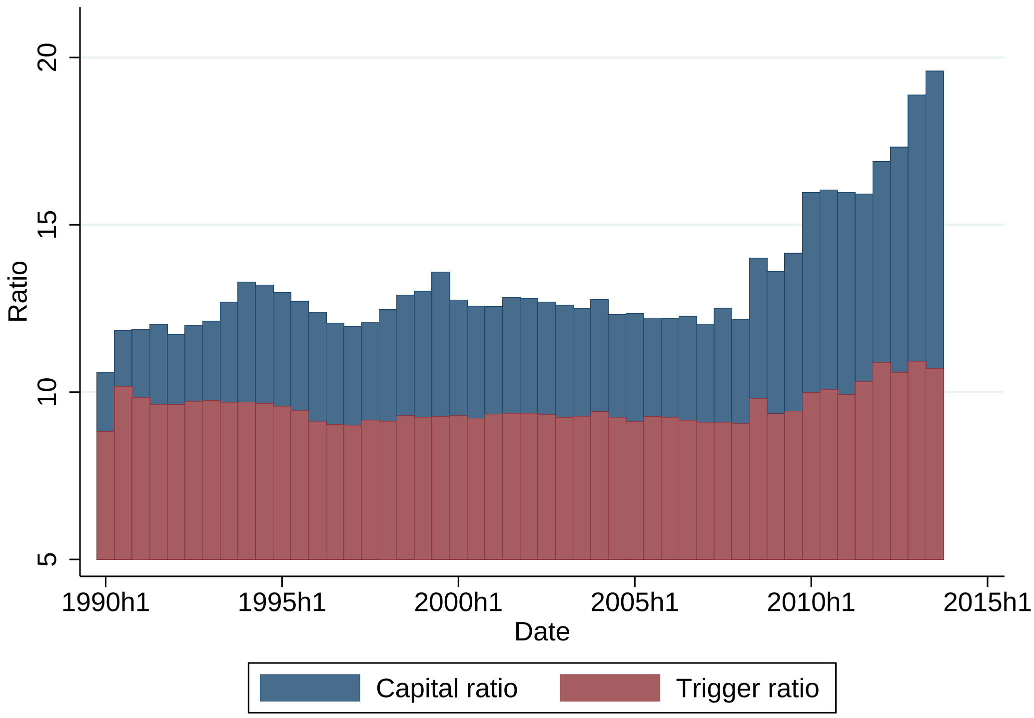
Note: this figure shows the distribution of half-year changes in trigger ratios over time. The points correspond to the lower adjacent value, 25$^{\text{th}}$ percentile, median, 75$^{\text{th}}$ percentile and upper adjacent value. I exclude values of trigger ratio changes with absolute values less than 0.1pps.

Figure 2.5: Box plot of capital buffers over time



Note: this figure shows the distribution of capital buffers over time. Capital buffers are calculated as the difference between a bank's risk-based capital ratio (total regulatory capital divided by total risk-weighted assets) and its trigger ratio. The points correspond to the lower adjacent value, 25$^{\text{th}}$ percentile, median, 75$^{\text{th}}$ percentile and upper adjacent value.

Figure 2.6: Time series of aggregate trigger and capital ratios



Note: this figure shows a time series of aggregate risk-based capital ratios and trigger ratios. Aggregate ratios are calculated as a weighted average of the individual bank ratios using total assets as weights.

6 Method

A priori, there are concerns with treating changes in trigger ratios as exogenous. One may expect changes in trigger ratios to be correlated with balance sheets risks. For example, a bank that undertakes riskier lending would be subject to greater credit risk, thus leading to different behavior compared to banks undertaking less risky lending. The increased credit risk may concern the regulator, leading to a higher capital requirement. There is thus a selection problem as it can be difficult to separately identify the causal effect of capital requirements from the potentially different nature of those banks receiving a tighter capital requirement. I therefore begin by empirically testing the assumption made in Aiyar et al. (2014) and Bridges et al. (2014) that changes in capital requirements are orthogonal to balance sheet risks. Using the least absolute shrinkage and selection operator (lasso) of Tibshirani (1996), I test whether key balance sheet variables enter into the regulator's reaction function. Upon verifying this assumption, I employ the local projection method of Jordá (2005) to trace out impulse responses of capital ratios and its subcomponents to a capital requirement change.

6.1 Lasso regressions

A major advantage of the HBRD dataset is that it contains a large amount of bank-level information that would have been observable to the regulator when setting requirements. I use lasso regressions to establish whether regulators consider balance sheet variables when setting banks' trigger ratios.[17] The standard lasso estimator minimizes the following objective function:

$$\arg\min_{\boldsymbol{\beta}} \frac{1}{N}\sum_{i,t}(y_{it} - \mathbf{x}'_{it}\boldsymbol{\beta})^2 + \frac{\lambda}{N}\|\boldsymbol{\beta}\|_1$$

where $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{J}|\beta_j|$, λ is the key penalization parameter and N is the number of observations used in the estimation. As such, the lasso regression seeks to minimize the residual sum of squares like in OLS estimation; however, unlike OLS it imposes an $\ell 1$-penalty on the coefficients. This penalty term shrinks the coefficients, some of which are shrunk down to zero, thus yielding sparse solutions and aiding model interpretation.[18] To obtain λ, I use K-fold cross-validation, which repeatedly partitions data into training and validation data and chooses the λ that minimizes an estimated mean squared prediction

[17] I use the *lassopack* Stata package of Ahrens et al. (2018) for lasso estimation.

[18] Alternative approaches to variable selection include stepwise techniques, best subset selection methods and least angle regression. There is no consensus over which approach should be preferred, particularly when there are a large number of explanatory variables (see Bertsimas et al. (2016) and Hastie et al. (2017) for further discussion).

error (MSPE).[19] Section B.2 describes the steps involved in the cross-validation procedure. I set $K = 10$ in the baseline analysis, but provide robustness checks using alternative values of K in Section 8.[20].

In my setting, y_{it} is the half-year change in trigger ratio, while $\mathbf{x}_{it}$ contains the one-period lag of annual growth rates or ratio changes of 30 bank balance sheet variables.[21] Table 2.3 provides the full list of variables.[22] The use of the one-period lag reflects lags associated with collating and communicating information about the bank to the regulator. I use annual rather than half-year movements in order to capture a general trend in bank behavior rather than higher frequency movements that could be driven by a temporary shock. Time dummies are included in $\mathbf{x}_{it}$ to capture sector-wide changes in capital requirements associated with, for example, macroeconomic fluctuations.[23]

6.2 Impulse responses following a capital requirement change

I apply local projections (Jordá, 2005) to trace out impulse responses of the capital ratio and its subcomponents. Under local projections, the model is estimated separately for each horizon h, thus allowing for flexibility in the shape of the impulse responses.[24] For each $h \in \{0, 1, ..., H\}$, I estimate the following model:

$$y_{i,t+h} - y_{i,t-1} = \beta_0^{(h)} + \beta_1^{(h)} \Delta trigger_{i,t} + \sum_{l=1}^{L} \delta_l^{(h)} \Delta y_{i,t-l} + \sum_{l=0}^{L} \boldsymbol{\eta}_l^{(h)} \Delta \mathbf{x}_{i,t-l} + \alpha_i^{(h)} + \gamma_t^{(h)} + \nu_{i,t}^{(h)} \tag{2.1}$$

where $y_{i,t+h}$ denotes the value of the variable of interest for bank i at time $t+h$, $\Delta trigger_{i,t}$ is the half-year change in trigger ratio, and $\mathbf{x}_{i,t-l}$ denotes a vector of controls for bank i in period $t-l$.[25] Lags of $\Delta y_{i,t-l}$ are included to sweep up serial correlation, and L denotes the maximal lag for the controls and the lags of $\Delta y_{i,t}$. $\alpha_i^{(h)}$ and $\gamma_t^{(h)}$ denote bank and time fixed effects respectively. I include the Tier 1 capital ratio and Tier 1 leverage ratio in the

[19] Alternative methods for selecting λ are discussed in Section 8.

[20] $K = 10$ is viewed to perform well on model selection (see Breiman and Spector, 1992; Kohavi, 1995; Zou (2006))

[21] A within transformation on all predictor variables ($\mathbf{x}$) is applied before doing the lasso regressions.

[22] As the lasso constraint involves the sum of absolute values of the $\boldsymbol{\beta}$ coefficients not exceeding some value, the variables in $\mathbf{x}_{it}$ are standardized to have zero mean and unit variance to ensure they are of the same scale. Note that only data in the training dataset is used when standardizing.

[23] Time dummies are partialled out prior to the lasso estimation in order to keep them in the final model. Partialling out a variable is equivalent to not penalizing that variable (Yamada, 2017).

[24] An advantage of local projections over vector autoregressions is that the former does not impose a dynamic structure, making it more robust to misspecification and less susceptible to the curse of dimensionality (Barnichon and Brownlees, 2016).

[25] If y is a quantity variable such as total loans or total regulatory capital, it is transformed into logs prior to estimation such that $y_{i,t+h} - y_{i,t-1}$ gives the cumulative growth from period $t-1$ to period $t+h$. No transformation is applied when y is a ratio.

vector of controls, $\mathbf{x}_{i,t-k}$, and set $L = 2$.

The impulse response of a variable y is given by plotting the estimates $\hat{\beta}_1^{(h)}$ over h. I take $H = 6$ such that the impulse responses look at the effect of capital requirements over a three-year period. For the estimates to be causal, I require that changes in trigger ratios are orthogonal to other bank- and time-varying drivers of the outcome of interest. The validity of this assumption is discussed in Section 7.1.

7 Results

7.1 Baseline regressions

In the baseline specification, I pool together all banks, thus assuming that the reaction functions of the regulator and the response of banks to a capital requirement change are common across banks. Table 2.2 shows the reaction function of the regulator following the lasso estimation. None of the 30 balance sheet variables included in the lasso are selected to feature in the reaction function. This finding is in line with the anecdotal evidence described in Section 3 which suggests that FSA regulators focused more on control risks such as IT systems rather than balance sheet risks. Together, this suggests that much of the variation in capital requirement changes comes from control risks, and so the assumption required for my estimates to be causal is for control risks to be orthogonal to balance sheet risks.

TABLE 2.2: Lasso-selected reaction function of the regulator (baseline specification)

	(1)
Constant	-0.00361
	(0.011)
Time fixed effects	Yes
Bank fixed effects	Yes
Observations	3256
Number of banking groups	212
R-squared	0.06

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses. This table shows the lasso estimation following Section 6.1. The dependent variable is the half-year change in trigger ratio. The variables that appear in the lasso estimation are given in Table 2.3. An intercept and bank and time fixed effects are included.

Figure 2.7 plots the impulse responses following a 1pp capital requirement increase under the baseline specification. The first key result is that there is a significant rise in capital ratios by just under 0.5pps immediately, showing that there is an instant, but partial, adjustment. The coefficient estimate remains stable and positive throughout. As far as three years later, capital ratios remain about 0.5pps larger than the pre-shock level. The next step is to understand which components of capital ratios adjust. Risk-based capital ratios can be decomposed as:

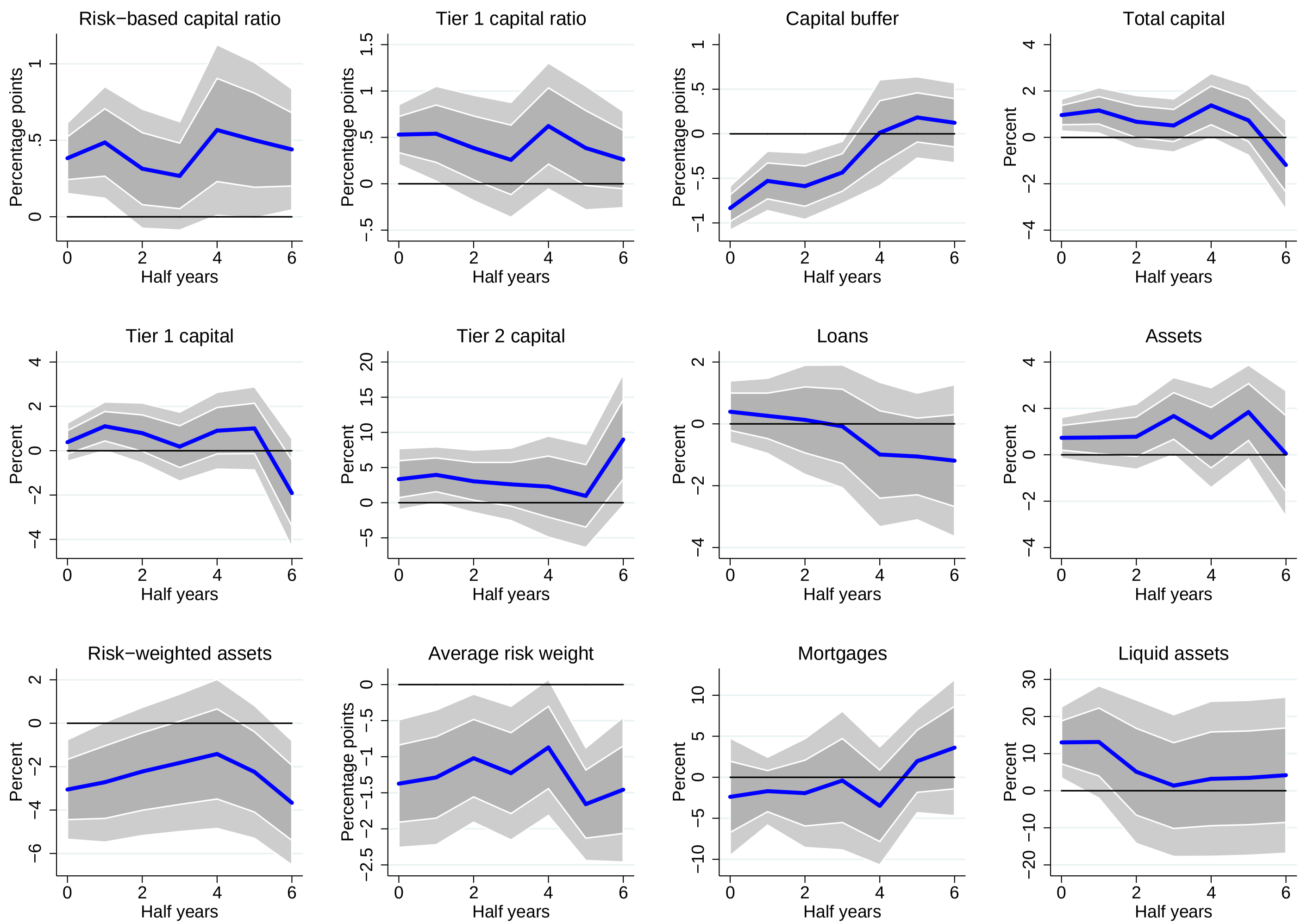
$$\begin{aligned}\text{Risk-based capital ratio} &= \frac{\text{Total regulatory capital}}{\text{Risk-weighted assets}} \\ &= \frac{\text{Total regulatory capital}}{\text{Total assets}} \times (\underbrace{\frac{\text{Risk-weighted assets}}{\text{Total assets}}}_{\text{Average risk weight}})^{-1}\end{aligned}$$

I can thus analyse whether the adjustment is mainly through a quantity effect, whereby the quantity of total regulatory capital and/or assets changes, or whether there is a shift in the risk composition of the bank's assets. The second key result is that there is a quantity effect for capital, but not for assets. From Figure 2.7, there is a significant increase in total regulatory capital of around 1% after one year. Tier 2 capital increases by just under 4% after a year compared to a rise of around 1% for Tier 1 capital. Given that Tier 2 capital is of lower quality and thus cheaper than Tier 1 capital, the decision of banks to use this type of capital makes sense from a cost-minimization perspective. Indeed, Francis and Osborne (2012) and de Ramon et al. (2022) also find that banks adjust through Tier 2 capital instruments. If anything, total assets rise following a capital requirement increase, which would lower capital ratios. The levels of loans and mortgages do not change significantly, suggesting that banks do not cut lending in response to a rise in capital requirements. Instead, liquid assets increase, which gives some evidence for a switch to less risky assets and can explain the rise in total assets. This risk composition effect is highlighted when looking at risk-weighted assets and average risk weights. The immediate-term reaction of risk-weighted assets is a decline of around 3%. Given that total assets, if anything, increase, there is a clear significant drop in average risk weights of around 1-1.5pps. As such, the third key finding is that the quantities of assets and loans do not fall in response to a rise in capital requirements; however, there is a composition effect towards less risky assets.

7.2 Heterogeneity analysis

In this section, I conduct micro-level heterogeneity analysis, redoing the lasso estimation and impulse responses for different subsamples based on time period (pre- vs. post-crisis), the direction of the capital requirement change and bank size. Summary statistics based

FIGURE 2.7: Impulse responses to a 1pp capital requirement increase



Note: this figure shows the local projection impulse responses from estimation of Equation 2.1 for twelve bank variables following a 1pp increase in capital requirements. Table 2.4 gives further details on the dependent variables. 68% and 90% confidence intervals are shown.

on these subsamples are given in Tables 2.5 to 2.7. First, I examine whether the reaction to a capital requirement change differs before and after the financial crisis.[26] As shown in Table 2.8, the lasso-selected model does not include any individual balance sheet variables. Figure 2.10 shows the impulse responses following a 1pp capital requirement increase. There are some notable differences in the two responses: first, the response of capital ratios is much more delayed in the pre-crisis period with risk-based capital ratios remaining unchanged until 18 months after the shock when there is a complete pass-through, meaning that banks completely dug into their capital buffers until then. In contrast, there is an immediate significant, but incomplete, pass-through into risk-based capital ratios in the post-crisis period. The swifter response could reflect greater supervisory intensity since the crisis. The second difference is that the pre-crisis response is particularly driven by capital accumulation, while the risk composition channel that was significant in the full sample results appears to be a post-crisis phenomenon. For the pre-crisis sample, the change in average risk weights is statistically insignificant for most horizons, while there is a significant drop during the first two years for the post-crisis period. The quantities of total assets and loans hardly change in the post-crisis period. However, there is a significant decline in total loans after two years for the pre-crisis period and so the quantity effect in assets plays more of a role here.

Next I look at whether banks behave differentially to a loosening versus a tightening of their capital requirements. I split changes in trigger ratios (after bank and time fixed effects have been stripped out) into positive and negative values. Figure 2.11 shows the impulse responses from this exercise. There are also differences between the two sets of responses here: first, capital ratios only adjust to declines in capital requirements. For an increase, banks respond by digging into their buffers. One concern with this result is that banks operating at their minimum requirement should adopt a complete pass-through of an increase in requirements. As such, the null overall effect would suggest that banks holding capital buffers have a negative pass-through, which would be unusual. However, there are very few banks operating with no or very small buffers. Figure 2.5 and Table 2.1 show that most banks do hold a buffer with the median buffer being almost 4.5pps. Even at the 10$^{\text{th}}$ percentile, the buffer is over 1pp, so almost all banks could absorb a 1pp rise in requirements through their buffers. While a further decomposition separating banks with very small buffers would be interesting, the limited number of such banks would make estimation imprecise. Second, the risk composition channel holds for declines in capital

[26] Pre-crisis observations are taken to be before and including 2007H1, and thus are all prior to the unravelling of interbank markets that arguably began with BNP Paribas stopping withdrawals from three investment funds on 9$^{\text{th}}$ August 2007.

requirements, but not increases. Following a loosening of requirements, the average risk weight increases, suggesting that banks move into riskier asset classes. However, other than in the immediate term, the average risk weight does not fall following a tightening of requirements.

I now look at whether small banks react differently to large banks to a capital requirement shock. I divide the sample based on the median value for total assets by each half-year period. Whilst such a distinction is naturally of interest, bank size is also an important determinant of the resource allocation of regulators as described in Section 3. For smaller banks, much of the risk assessment is through baseline monitoring of regulatory returns, while for larger banks, factors such as on-site visits also play a role. As a result, it is possible that balance sheet risks identified using regulatory returns play a larger role when deciding upon small banks' capital requirements. Table 2.9 shows the lasso-selected model, but again gives the result that no individual balance sheet variables are selected, even for small banks.[27] Figure 2.12 shows that much of the reaction to capital requirement changes comes from small banks. I find that for small banks, risk-based capital ratios rise by around 0.5pps after one year, while for large banks, there is no adjustment of risk-based capital ratios during the first 18 months. As a result, there is a larger depletion of capital buffers for large banks during this phase. The reaction of small banks appears to come through a combination of two channels, namely a quantity effect for capital and a risk composition effect towards less risky assets, but not a quantity effect for total assets or loans. The significant rise in total capital of over 1% seems to be predominantly through Tier 2 rather than Tier 1 capital, although the latter does show a significant increase too. For large banks, there is no significant change in total capital until two years after the regulatory change, which is in line with the behaviour of risk-based capital ratios. Total assets, if anything, increase over time, meaning there is no quantity effect in total assets. There is a slight decline in average risk weights, but this is much smaller than for small banks. As such, compared with small banks, the reaction of large banks is much more delayed and comes predominantly through capital accumulation with a small risk composition effect. However, as shown in the summary statistics by size of bank in Table 2.7, there are significant differences between the two subgroups other than total assets. For example, large banks seem to have smaller capital buffers, loan-to-assets ratios and solvency ratios, all of which may interact with the response to a capital requirement shock. As such, the results for small versus large banks should be taken lightly, as it will require a larger sample size to split the subgroup further in order to isolate the impact of bank

[27] This result is robust to using the 75^{th} and 90^{th} percentiles for total assets to separate small and large banks.

size with reasonable precision.

8 Robustness

One concern with the methodology used in this paper is whether lasso techniques are appropriate for model selection. A necessary and sufficient condition for consistent variable selection is for the *irrepresentable condition* to be satisfied (Meinshausen and Bühlmann, 2006; Zhao and Yu, 2006; Zou, 2006). This condition requires the correlation between variables inside the regulator's actual reaction function and variables outside of the true model to be sufficiently low.[28] To explain intuitively why this condition is needed, suppose that there are two highly correlated variables, but only one enters into the true reaction function. The lasso procedure may then select the other variable as a result of the high correlation, leading to incorrect conclusions for variable selection. This would be irrespective of the sample size or the degree of regularization. In my setting, it is arguably difficult to satisfy the irrepresentable condition as decisions of banks across a wide range of balance sheet variables are likely to be correlated. I therefore apply the adaptive lasso of Zou (2006), which is consistent for variable selection under weaker assumptions. The adaptive lasso involves a two-step procedure: first, an initial estimator $\hat{\boldsymbol{\beta}}_{\text{initial}}$ is obtained using a standard fixed effects regression:

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i + \gamma_t + \epsilon_{it}$$

where y_{it} is the half-year change in trigger ratio for bank i at time t and $\boldsymbol{x_{it}}$ is the vector of all 30 balance sheet variables. The adaptive lasso estimator is then:

$$\hat{\boldsymbol{\beta}}_{\text{adaptive}} = \arg\min_{\boldsymbol{\beta}} \frac{1}{N}\sum_{i,t}(y_{it} - \mathbf{x}_{it}'\boldsymbol{\beta})^2 + \frac{\lambda}{N}\sum_{j=1}^{p}\frac{|\beta_j|}{|\hat{\beta}_{\text{initial},j}|^{\theta}}$$

[28]More formally, the irrepresentable condition is as follows: denote $\hat{\Sigma} \equiv n^{-1}\mathbf{X}^T\mathbf{X}$ and define $S_0 = \{j : \beta_j \neq 0\}$ as the set of variables that do belong in the true model. Without loss of generality, suppose $S_0 = \{1, 2, ..., s_0\}$, so the set S_0 contains the first s_0 variables. Writing in block form:

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{1,1} & \hat{\Sigma}_{1,2} \\ \hat{\Sigma}_{2,1} & \hat{\Sigma}_{2,2} \end{pmatrix}$$

where $\hat{\Sigma}_{1,1}$ is an $s_0 \times s_0$ matrix for those variables in S_0, $\hat{\Sigma}_{1,2} = \hat{\Sigma}_{2,1}^T$ is an $s_0 \times (p - s_0)$ matrix (where p is the total number of variables) and $\hat{\Sigma}_{2,2}$ is a $(p - s_0) \times (p - s_0)$ matrix. The irrepresentable condition states:

$$|\hat{\Sigma}_{2,1}\hat{\Sigma}_{1,1}^{-1}\text{sign}(\beta_1, ..., \beta_{s_0})| \leq \boldsymbol{\theta}$$

where $\boldsymbol{\theta}$ is a $(p - s_0) \times 1$ column vector with each element $0 < \theta_j < 1$ and $\text{sign}(\beta_1, ..., \beta_{s_0}) = (\text{sign}(\beta_1), ..., \text{sign}(\beta_p))^T$. The inequality must hold element-wise.

where as in Section 6.1, a within transformation is applied to the regressors $\mathbf{x}$ and time fixed effects are partialled out prior to estimation. λ is again obtained through cross validation and I take $\theta = 1$. Column 1 of Table 2.10 shows that the adaptive lasso also selects no variables.

I also consider other lasso variants. As an alternative to cross validation for the selection of the penalization parameter λ, I also use the clustered rigorous lasso of Belloni et al. (2016), which provides a theoretically-driven and data-dependent penalization:

$$\lambda = 2c\sqrt{N}\Phi^{-1}(1 - \frac{\gamma}{2p})$$

where γ is the number of clusters (i.e. the number of bank groups in the sample) and $c = 1.1$. p is the number of penalized variables in the lasso (in my case, 30) and Φ^{-1} is the inverse normal CDF function. The results from this are given in Column 2 of Table 2.10. Column 3 gives the model selected through the Extended Bayesian Information Criteria (EBIC) of Chen and Chen (2008). In both cases, the lasso procedure again selects no variables. Columns 4-7 consider K-fold cross validation for different values of K (in the baseline specification, I use $K = 10$). In all cases, no variables are chosen.

A further concern is that the full sample contains observations where a supervisory review may not have taken place in that period. Including these observations could make it difficult to identify the regulator's reaction function. As described in Section 3, supervisory reviews tend to occur at fixed time intervals of every 1-3 years. If the supervisory review dates were recorded, one could simply exclude observations where a review recently took place or focus on observations when a review did occur. However, these dates are not recorded in the dataset. I provide two proxy approaches based on observed changes in capital requirements. Column 8 shows the selected model under the baseline approach using only those observations for which a change actually took place. Here no variables are selected. Column 9 considers only those observations where there are no observed changes in trigger ratios in the previous 12 months given the supervisory cycle length of 1-3 years. In this case, more balance sheet variables are selected. In particular, lower risk-weighted asset growth and higher unsecured loans and non-performing loans growth rates are associated with higher capital requirements. Figures 2.13 and 2.14 show the impulse responses for these two samples.[29] The main findings from the baseline specification of Figure 2.7 appear to hold in both cases, although a medium-term quantity effect for loans

[29]For the specification with no changes in trigger ratios in the past 12 months, changes in capital ratios are stripped of the lasso-selected balance sheet variables and bank and time fixed effects, and the residuals are used in the impulse responses.

appears to occur when using observations where no requirement change occurred in the preceding 12 months.

9 Conclusion

A growing interest in micro- and macroprudential regulation and their impacts on the banking sector has emerged since the financial crisis. This paper seeks to quantify the effect of one particular tool, namely bank capital requirements. Using confidential data on individual bank capital requirements in the UK from 1989H1-2013H2, I study the impact of changes in individual bank capital requirements on bank balance sheet behavior. I first show using lasso techniques that changes in capital requirements appear orthogonal to balance sheet risks. Using local projections à la Jordá (2005), I then trace out the impulse responses of capital ratios and its subcomponents over a three-year window following a capital requirements change.

Using the full sample of banks, I find that, on average, banks do adjust their actual capital ratios following a change in requirements, but only partially with a pass-through of around 50%. Much of this reaction comes through capital accumulation, in particular the level of Tier 2 capital; however, the quantity of loans is unchanged. There is also evidence of a composition effect, whereby banks adjust the average riskiness of their asset portfolio. Banks only react to decreases in capital requirements with an increase in requirements being absorbed by banks' pre-existing capital buffers. Comparing the impact of a capital requirement change in the pre- and post-financial crisis periods, I find that the pre-crisis response is characterized by quantity effects in capital and loans, but no composition effect. In particular, total lending falls by 5% on average one year after a 1pp increase in capital requirements. Instead, the post-crisis period is associated with no significant change in the quantity of loans, though there is a composition effect towards less risky assets.

References

AHRENS, A., C. B. HANSEN, AND M. E. SCHAFFER (2018): "LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation," Statistical Software Components, Boston College Department of Economics.

AIYAR, S., C. W. CALOMIRIS, AND T. WIELADEK (2014): "Does Macro-Prudential Regulation Leak? Evidence from a UK Policy Experiment," *Journal of Money, Credit and Banking*, 46, 181–214.

ALFON, I., I. ARGIMON, AND P. BASCUÑANA-AMBRÓS (2004): "What determines how much capital is being held by UK banks and building societies?," Discussion Paper 22, Financial Services Authority.

AYUSO, J., D. PÉREZ, AND J. SAURINA (2004): "Are capital buffers pro-cyclical?: Evidence from Spanish panel data," *Journal of Financial Intermediation*, 13(2), 249–264.

BANK OF ENGLAND (2018): "The Prudential Regulation Authority's approach to banking supervision," Discussion paper, Bank of England.

BARNICHON, R., AND C. BROWNLEES (2016): "Impulse Response Estimation By Smooth Local Projections," Discussion Paper 11726, C.E.P.R. Discussion Papers.

BASEL COMMITTEE ON BANKING SUPERVISION (1988): "International convergence of capital measurement and capital standards," Discussion paper, Bank of International Settlements.

——— (2006): "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version," Discussion paper, Bank of International Settlements.

——— (2009): "Strengthening the resilience of the banking sector," Discussion paper, Bank of International Settlements.

——— (2010a): "Basel III: A global regulatory framework for more resilient banks and banking systems," Discussion paper, Bank of International Settlements.

——— (2010b): "Basel III: International framework for liquidity risk measurement, standards and monitoring," Discussion paper, Bank of International Settlements.

Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016): "Inference in High-Dimensional Panel Models With an Application to Gun Control," *Journal of Business & Economic Statistics*, 34(4), 590–605.

Bernanke, B. S., C. S. Lown, and B. M. Friedman (1991): "The Credit Crunch," *Brookings Papers on Economic Activity*, 1991(2), 205–247.

Bertsimas, D., A. King, and R. Mazumder (2016): "Best subset selection via a modern optimization lens," *The Annals of Statistics*, 44(2), 813–852.

Breiman, L., and P. Spector (1992): "Submodel Selection and Evaluation in Regression. The X-Random Case," *International Statistical Review*, 60(3), 291–319.

Bridges, J., D. Gregory, M. Nielsen, S. Pezzini, A. Radia, and M. Spaltro (2014): "The impact of capital requirements on bank lending," *Bank of England Staff Working Papers*, 486.

Chen, J., and Z. Chen (2008): "Extended Bayesian information criteria for model selection with large model spaces," *Biometrika*, 95, 759–771.

De Jonghe, O., H. D. Mulier, K. Mulier, S. Ongena, and G. Schepens (2019): "Some Borrowers Are More Equal than Others: Bank Funding Shocks and Credit Reallocation," *Review of Finance*.

de Ramon, S. J., W. B. Francis, and Q. Harris (2022): "Bank-specific capital requirements and capital management from 1989-2013: Further evidence from the UK," *Journal of Banking & Finance*, 138, 106189.

de-Ramon, S. J. A., W. Francis, and K. Milonas (2017): "An overview of the UK banking sector since the Basel Accord: insights from a new regulatory database," *Bank of England Staff Working Papers*, 652.

Diamond, D. W. (1984): "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies*, 51(3), 393–414.

Ediz, T., I. Michael, and W. Perraudin (1998): "The impact of capital requirements on U.K. bank behaviour," *FRBNY Economic Policy Review*, pp. 15–22.

Elliott, D. J. (2009): "Quantifying the effects on lending of increased capital requirements," *The Brookings Institution*.

Fang, X., D. Jutrsa, M. S. M. Peria, A. F. Presbitero, L. Ratnovski, and F. J. Vardy (2018): "The Effects of Higher Bank Capital Requirements on Credit in Peru," *IMF Working Papers*, 18/222.

Financial Services Authority (1998): "Risk based approach to supervision of banks," Discussion paper, Financial Services Authority.

——— (2000): "Building the new regulator: Progress report 1," Discussion paper, Financial Services Authority.

——— (2002): "Building the new regulator: Progress report 2," Discussion paper, Financial Services Authority.

——— (2008): "The supervision of Northern Rock: a lessons learned review," Discussion paper, Financial Services Authority.

——— (2009): "The Turner Review: A regulatory response to the global banking crisis," Discussion paper, Financial Services Authority.

Fonseca, A. R., and F. González (2010): "How bank capital buffers vary across countries: The influence of cost of deposits, market power and bank regulation," *Journal of Banking & Finance*, 34(4), 892–902.

Francis, W. B., and M. Osborne (2012): "Capital requirements and bank behavior in the UK: Are there lessons for international capital standards?," *Journal of Banking & Finance*, 36(3), 803–816.

Gale, D., and M. Hellwig (1985): "Incentive-Compatible Debt Contracts: The One-Period Problem," *The Review of Economic Studies*, 52, 647–663.

Gropp, R., T. Mosk, S. Ongena, and C. Wix (2018): "Banks Response to Higher Capital Requirements: Evidence from a Quasi-Natural Experiment," *The Review of Financial Studies*, 32, 266–299.

Hancock, D., and J. A. Wilcox (1993): "Has There Been a "Capital Crunch" in Banking? The Effects on Bank Lending of Real Estate Market Conditions and Bank Capital Shortfalls," *Journal of Housing Economics*, 3(1), 31–50.

Hancock, D., and J. A. Wilcox (1994): "Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations," *Real Estate Economics*, 22, 59–94.

Hanson, S. G., A. K. Kashyap, and J. C. Stein (2011): "A Macroprudential Approach to Financial Regulation," *Journal of Economic Perspectives*, 25(1), 3–28.

Hastie, T., R. Tibshirani, and R. J. Tibshirani (2017): "Extended comparisons of best subset selection, forward stepwise selection, and the lasso," *arXiv preprint arXiv:1707.08692*.

Heid, F., D. Porath, and S. Stolz (2004): "Does capital regulation matter for bank behaviour? Evidence for German savings banks," Discussion paper series 2: Banking and financial studies, Deutsche Bundesbank.

International Monetary Fund (2003): "United Kingdom: Financial Sector Assessment Program Technical Notes and Detailed Standards Assessments," Discussion Paper 03/208, International Monetary Fund.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina (2010): "Credit supply: identifying balance-sheet channels with loan applications and granted loans," Working Paper 1030, ECB Working Paper series.

——— (2017): "Macroprudential Policy, Countercyclical Bank Capital Buffers, and Credit Supply: Evidence from the Spanish Dynamic Provisioning Experiments," *Journal of Political Economy*, 125(6), 2126–2177.

Jokipii, T., and A. Milne (2008): "The cyclical behaviour of European bank capital buffers," *Journal of Banking & Finance*, 32(8), 1440–1451.

Jordá, Ò. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95(1), 161–182.

Kashyap, A. K., J. C. Stein, and S. Hanson (2010): "An Analysis of the Impact of "substantially heightened" Capital Requirements on Large Financial Institutions," *Mimeo*.

Kohavi, R. (1995): "A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, pp. 1137–1143. Morgan Kaufmann Publishers Inc.

Lindquist, K.-G. (2004): "Banks' buffer capital: how important is risk," *Journal of International Money and Finance*, 23(3), 493–513, Banking, Development and Structural Change.

Meeks, R. (2017): "Capital regulation and the macroeconomy: Empirical evidence and macroprudential policy," *European Economic Review*, 95, 125–141.

Meinshausen, N., and P. Bühlmann (2006): "High-Dimensional Graphs and Variable Selection with the Lasso," *The Annals of Statistics*, 34(3), 1436–1462.

Mésonnier, J.-S., and A. Monks (2015): "Did the EBA Capital Exercise Cause a Credit Crunch in the Euro Area?," *International Journal of Central Banking*, 11(3), 75–117.

Miles, D., J. Yang, and G. Marcheggiano (2013): "Optimal Bank Capital," *The Economic Journal*, 123, 1–37.

Modigliani, F., and M. H. Miller (1958): "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review*, 48(3), 261–297.

Montagnoli, A., K. Mouratidis, and K. Whyte (2018): "Assessing the Cyclical Behaviour of Bank Capital Buyers in a Finance-Augmented Macro-Economy," Working Papers 2018003, The University of Sheffield, Department of Economics.

Myers, S. C., and N. S. Majluf (1984): "Corporate financing and investment decisions when firms have information that investors do not have," *Journal of Financial Economics*, 13(2), 187–221.

Peek, J., and E. S. Rosengren (1997): "The International Transmission of Financial Shocks: The Case of Japan," *American Economic Review*, 87(4), 495–505.

Peura, S., and J. Keppo (2006): "Optimal Bank Capital with Costly Recapitalization," *The Journal of Business*, 79(4), 2163–2202.

Shim, J. (2013): "Bank capital buffer and portfolio risk: The influence of business cycle and revenue diversification," *Journal of Banking & Finance*, 37(3), 761–772.

Sironi, A. (2018): "The Evolution of Banking Regulation Since the Financial Crisis: A Critical Assessment," *BAFFI CAREFIN Centre Research Paper*, 2018-103.

Stolz, S., and M. Wedow (2011): "Banks' regulatory capital buffer and the business cycle: Evidence for Germany," *Journal of Financial Stability*, 7(2), 98–110.

Tibshirani, R. (1996): "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288.

Valencia, O. C., and A. O. Bolaños (2018): "Bank capital buffers around the world: Cyclical patterns and the effect of market power," *Journal of Financial Stability*, 38, 119–131.

VanHoose, D. (2007): "Theories of bank behavior under capital regulation," *Journal of Banking & Finance*, 31(12), 3680–3697.

——— (2008): "Bank capital regulation, economic stability and monetary policy: what does the academic literature tell us?," *Atlantic Economic Journal*, 36, 1–14.

Yamada, H. (2017): "The Frisch-Waugh-Lovell theorem for the lasso and the ridge regression," *Communications in Statistics - Theory and Methods*, 46(21), 10897–10902.

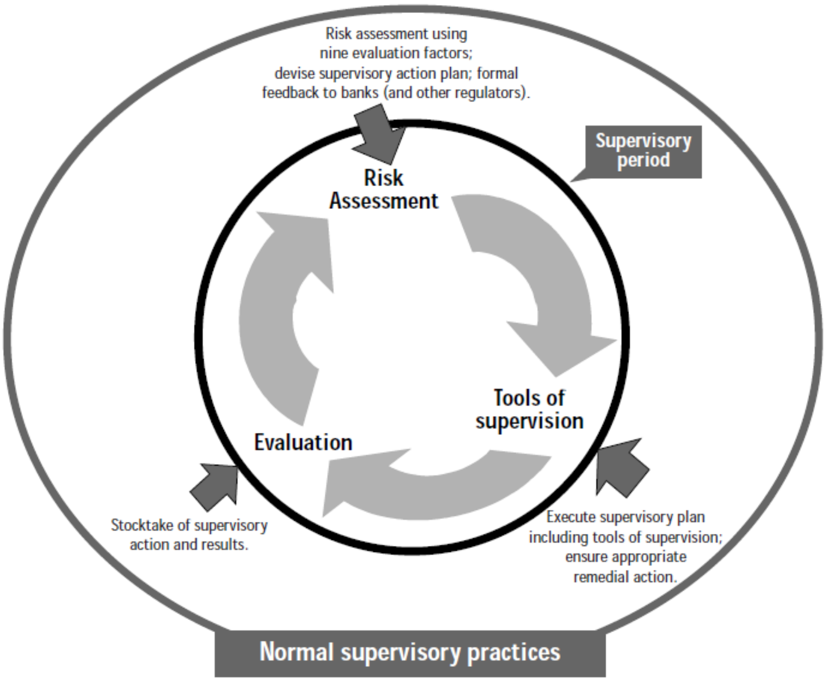
Zhao, P., and B. Yu (2006): "On Model Selection Consistency of Lasso," *Journal of Machine Learning Research*, 7, 2541–2563.

Zou, H. (2006): "The Adaptive Lasso and Its Oracle Properties," *Journal of the American Statistical Association*, 101(476), 1418–1429.

Appendix

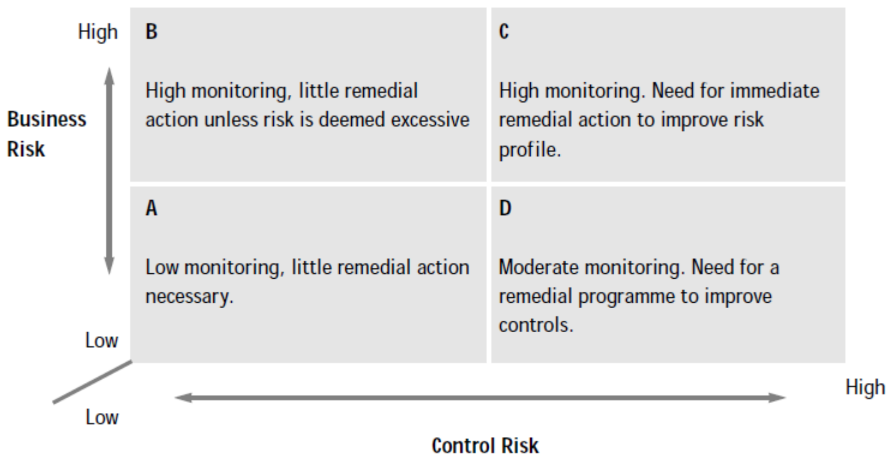
A Supervisory frameworks

Figure 2.8: RATE framework process



Note: this figure provides an overview of the FSA's RATE (Risk Assessment, Tools and Evaluation) framework, and is taken from Financial Services Authority (1998). Further details of the UK supervisory frameworks are provided in Section 3.

Figure 2.9: Supervisory intensity under the FSA's RATE framework



Note: this matrix gives a summary of the likely supervisory intensity following different combinations of business and control risk profiles. Figure is taken from Financial Services Authority (1998). Further details of the UK supervisory frameworks are provided in Section 3.

B Methodology details

B.1 Data cleaning steps

1. There are missing values in the dataset. These can either be specific variables missing in the otherwise-completed returns, or no returns at all for that bank in the given half year.[30] I linearly interpolate the data whenever there is a missing value for a variable, but data is available for the two periods on either side of that date.

2. Different banks file returns at different times in the year with the convention being June and December reporting. The varying length of one period for each bank makes it difficult to analyse the impact of capital requirements over different horizons. I use linear interpolation to align reporting period ends to the June/December convention such that one period corresponds to six months for all banks.

3. I replace suspicious zeros with missing values and use absolute values when a negative number is reported for a variable that should only permit weakly positive values.

4. I treat the banking group resulting from mergers and acquisitions as a new banking group as in de Ramon et al. (2022). Due to different financial structures, business strategies and management following such activity, it would not be appropriate to treat the resulting banking group as the same entity. In addition, I create a new institution whenever the half-year growth of total assets, loans or regulatory capital exceeds 50% in order to capture structural changes not covered by the identification of mergers in HBRD.[31]

5. To mitigate the impact of outliers, I drop observations where the half-year growth of total regulatory capital, assets or loans exceeds 50%, or if the half-year change in the trigger ratio exceeds 10pps. I also winsorize all variables at the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of a given half year.

B.2 Steps for K-fold cross-validation

The steps are as follows:

1. Data is partitioned into K folds of roughly equal size.

[30] In cases where entire reports are missing, this is typically due to special waivers being granted or because the regulator did not supervise the bank until a later period.

[31] An example of such a change is the merger of NatWest with Royal Bank of Scotland in 2000. A similar approach is also taken in de Ramon et al. (2022).

2. The first fold becomes the validation dataset and the other $K-1$ folds make the training dataset. For a given λ, the model is fit to the training data. Denoting the coefficient estimates from this step as $\hat{\beta}_{1,\lambda}$, the MSPE for fold 1 is:

$$MSPE_{1,\lambda} = \frac{1}{n_1}\sum_{i,t}^{n_1}(y_{it} - \mathbf{x}_{it}'\hat{\boldsymbol{\beta}}_{1,\lambda})$$

   Note that you sum over only those observations belonging to the validation dataset, which in this case are observations in fold 1.

3. Repeat the process using different folds as the validation dataset and compute $MSPE_{k,\lambda}$ for $k = 2, 3, ..., K$.

4. For a given λ, the K-fold cross-validation estimate of the MSPE, CV_λ, provides a measure of prediction performance and is computed as:

$$CV_\lambda = \frac{1}{K}\sum_{k=1}^{K} MSPE_{k,\lambda}$$

5. Repeat the above steps for multiple values of λ. The chosen λ, denoted as λ^*, is then:

$$\lambda^* = \arg\min_{\lambda} CV_\lambda$$

C Data and results

TABLE 2.3: Variables used in lasso regressions

Variable	Formula	Notes
Change in losses to loans ratio	$100 \times \Delta_2(\frac{\text{Write offs net of recoveries}}{\text{Loans}})$	Seasonally-adjusted value of net write-offs used.
Change in provisions to loans ratio	$100 \times \Delta_2(\frac{\text{Provisions}}{\text{Loans}})$	Total provisions includes specific and general provisions against bad or doubtful debt.
Change in impairments to assets ratio	$100 \times \Delta_2(\frac{\text{Impairments charge}}{\text{Assets}})$	Impairments charge is seasonally adjusted and includes the net charge or credit to the P&L account for the provision for doubtful debts.
Change in average risk weight	$100 \times \Delta_2(\frac{\text{RWA}}{\text{Assets}})$	
Change in loans to assets ratio	$100 \times \Delta_2(\frac{\text{Loans}}{\text{Assets}})$	Total loans includes all funds lent to counterparties other than credit institutions, central governments and central banks.
Change in loans to deposits ratio	$100 \times \Delta_2(\frac{\text{Loans}}{\text{Deposits}})$	Total deposits covers all intra-financial and retail deposits.
Change in liquid asset ratio	$100 \times \Delta_2(\frac{\text{High quality liquid assets}}{\text{Assets}})$	High quality liquid assets cover cash and balances at central banks, gilts, Treasury bills and other highly liquid bills.
Change in Tier 1 leverage ratio	$100 \times \Delta_2(\frac{\text{Tier 1 capital}}{\text{Assets}})$	
Change in solvency ratio	$100 \times \Delta_2(\frac{\text{Capital}}{\text{Required capital}})$	Capital is total regulatory capital held by the bank. Total required capital is given by the trigger ratio multiplied by total risk-weighted assets.

Change in efficiency ratio	$100 \times \Delta_2(\frac{\text{Overhead costs}}{\text{Non-interest income}})$	Total overhead costs include staff expenses, administrative costs and other operating expenses. Total non-interest income includes net-interest income, fee and commission income, other operating income and trading income.
Change in residential loans to assets ratio	$100 \times \Delta_2(\frac{\text{Residential}}{\text{Assets}})$	Total residential loans are all loans secured on residential property.
Change in capital buffer	$100 \times \Delta_2(\frac{\text{Capital}-\text{Required capital}}{\text{Required capital}})$	
Change in capital ratio	$100 \times \Delta_2(\frac{\text{Capital}}{\text{Assets}})$	
Change in Core Tier 1 capital ratio	$100 \times \Delta_2(\frac{\text{Core Tier 1 capital}}{\text{Assets}})$	Core Tier 1 capital includes all permanent share capital, reserves, share premium account, externally-verified interim net profits but excludes intangible assets and investments in own shares.
Change in non-core Tier 1 capital ratio	$100 \times \Delta_2(\frac{\text{Non Core Tier 1 capital}}{\text{Assets}})$	
Change in earning assets to total assets ratio	$100 \times \Delta_2(\frac{\text{Earning assets}}{\text{Total assets}})$	Earning assets are total assets net of cash & balances at central banks, intangible assets and fixed assets.
Change in interest income to earning assets ratio	$100 \times \Delta_2(\frac{\text{Interest income}}{\text{Earning assets}})$	Interest income includes income from interest received and accrued interest that has not yet been collected.
Change in interest expense to earning assets	$100 \times \Delta_2(\frac{\text{Interest expense}}{\text{Earning assets}})$	Interest expense includes interest paid and interest payable that has been accrued, but has not been collected yet.
Change in provisions ratio	$100 \times \Delta_2(\frac{\text{Impairments charge}_{i,t}}{\frac{1}{2}(\text{Assets}_{i,t-1}+\text{Assets}_{i,t-2})})$	

Change in net-interest income to earning assets	$100 \times \Delta_2(\frac{\text{Net interest income}}{\text{Earning assets}})$	Net interest income is the difference between interest income and interest expense.
Change in net operating income ratio	$100 \times \Delta_2(\frac{\text{Post-tax net income}_{i,t}}{\frac{1}{2}(\text{Assets}_{i,t-1}+\text{Assets}_{i,t-2})})$	Post-tax net income is total profits for the financial year up to the reporting date.
Assets growth	$100 \times \frac{\text{Assets}_{i,t}-\text{Assets}_{i,t-2}}{\text{Assets}_{i,t-2}}$	
Tier 1 capital growth	$100 \times \frac{\text{Tier 1 capital}_{i,t}-\text{Tier 1 capital}_{i,t-2}}{\text{Tier 1 capital}_{i,t-2}}$	
Total capital growth	$100 \times \frac{\text{Capital}_{i,t}-\text{Capital}_{i,t-2}}{\text{Capital}_{i,t-2}}$	
Loans growth	$100 \times \frac{\text{Loans}_{i,t}-\text{Loans}_{i,t-2}}{\text{Loans}_{i,t-2}}$	
Deposits growth	$100 \times \frac{\text{Deposits}_{i,t}-\text{Deposits}_{i,t-2}}{\text{Deposits}_{i,t-2}}$	
Risk-weighted assets growth	$100 \times \frac{\text{RWA}_{i,t}-\text{RWA}_{i,t-2}}{\text{RWA}_{i,t-2}}$	
Unsecured lending growth	$100 \times \frac{\text{Unsecured}_{i,t}-\text{Unsecured}_{i,t-2}}{\text{Unsecured}_{i,t-2}}$	Unsecured loans covers all funds lent to counterparties other than credit institutions excluding loans fully secured on residential property.
Residential loans growth	$100 \times \frac{\text{Residential}_{i,t}-\text{Residential}_{i,t-2}}{\text{Residential}_{i,t-2}}$	
Non-performing loans growth	$100 \times \frac{\text{Non-perform}_{i,t}-\text{Non-perform}_{i,t-2}}{\text{Non-perform}_{i,t-2}}$	

Note: this table provides a list of all variables considered in the lasso regressions of Section 6.1. Δ_2 refers to the annual change in the variable x. Growth rates are all annual. Data is from the Bank of England's HBRD dataset.

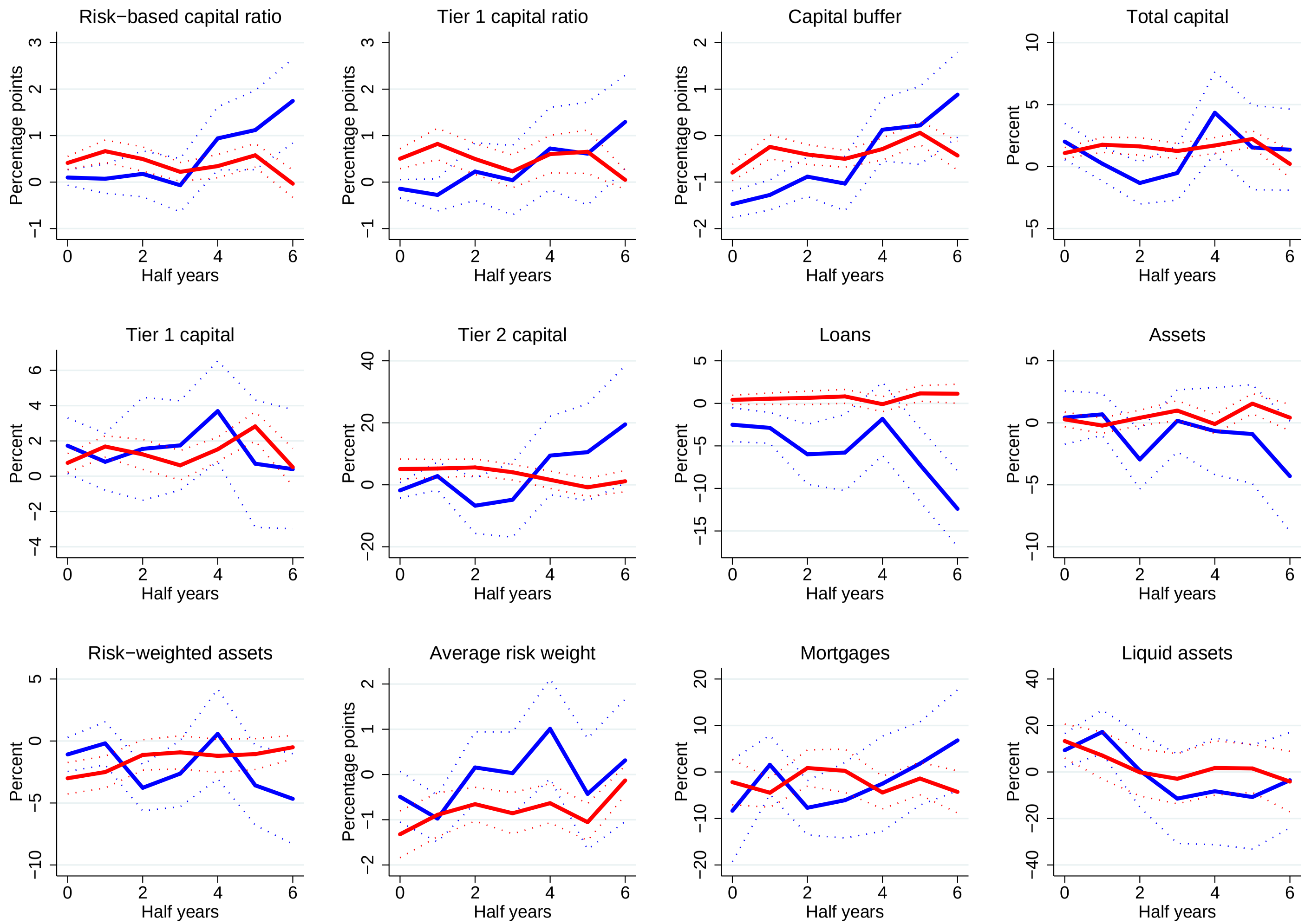
Table 2.4: Dependent variables used in micro-level analysis

Variable	**Formula**	**Notes**
Risk-based capital ratio	$100 \times \frac{\text{Capital}}{\text{RWA}}$	RWA denotes risk-weighted assets.
Tier 1 capital ratio	$100 \times \frac{\text{Tier 1 capital}}{\text{RWA}}$	
Capital buffer	$100 \times \frac{\text{Capital}-\text{Required capital}}{\text{Required capital}}$	Total required capital is given by the trigger ratio multiplied by total risk-weighted assets.
Average risk weight	$100 \times \frac{\text{RWA}}{\text{Assets}}$	

Liquid assets	Liquid assets here cover highly liquid assets as well as intra-financial deposits and other debt securities. High quality liquid assets are cash and balances at central banks, gilts, Treasury bills and other highly liquid bills.

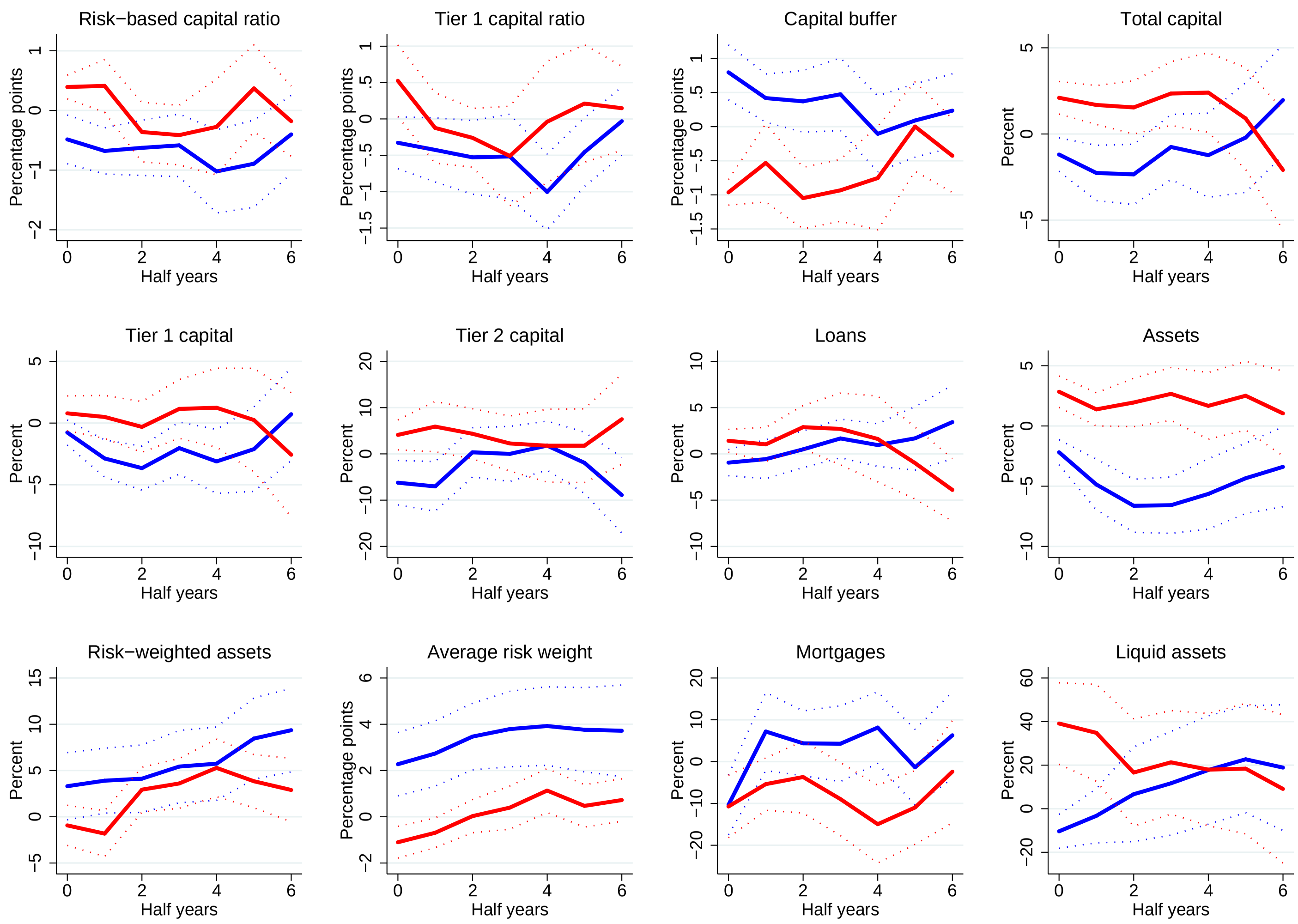
Note: this table provides details on the dependent variables used in the local projections of Equation 2.1. All data is from the Bank of England's HBRD dataset.

FIGURE 2.10: Impulse responses to a 1pp capital requirement increase: pre- vs. post-crisis



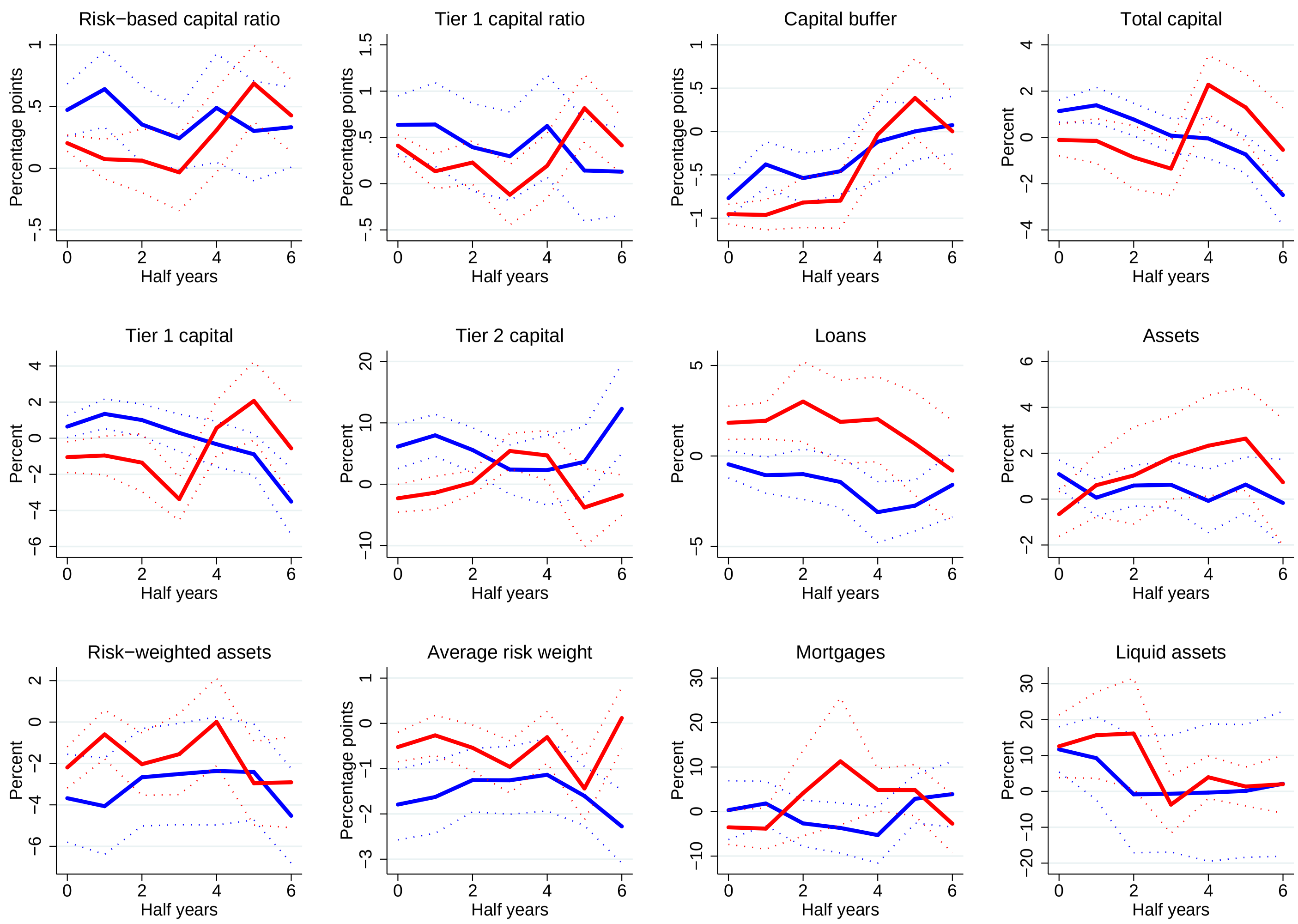
Note: this figure shows the local projection impulse responses for the pre-crisis (in blue) and post-crisis (in red) periods separately following estimation of Equation 2.1. The impulse responses are for a 1pp increase in capital requirements. Pre-crisis covers dates up to and including 2007H1. Table 2.4 gives further details on the dependent variables. 68% confidence intervals are shown.

FIGURE 2.11: Impulse responses to a 1pp capital requirement increase vs. a 1pp decrease



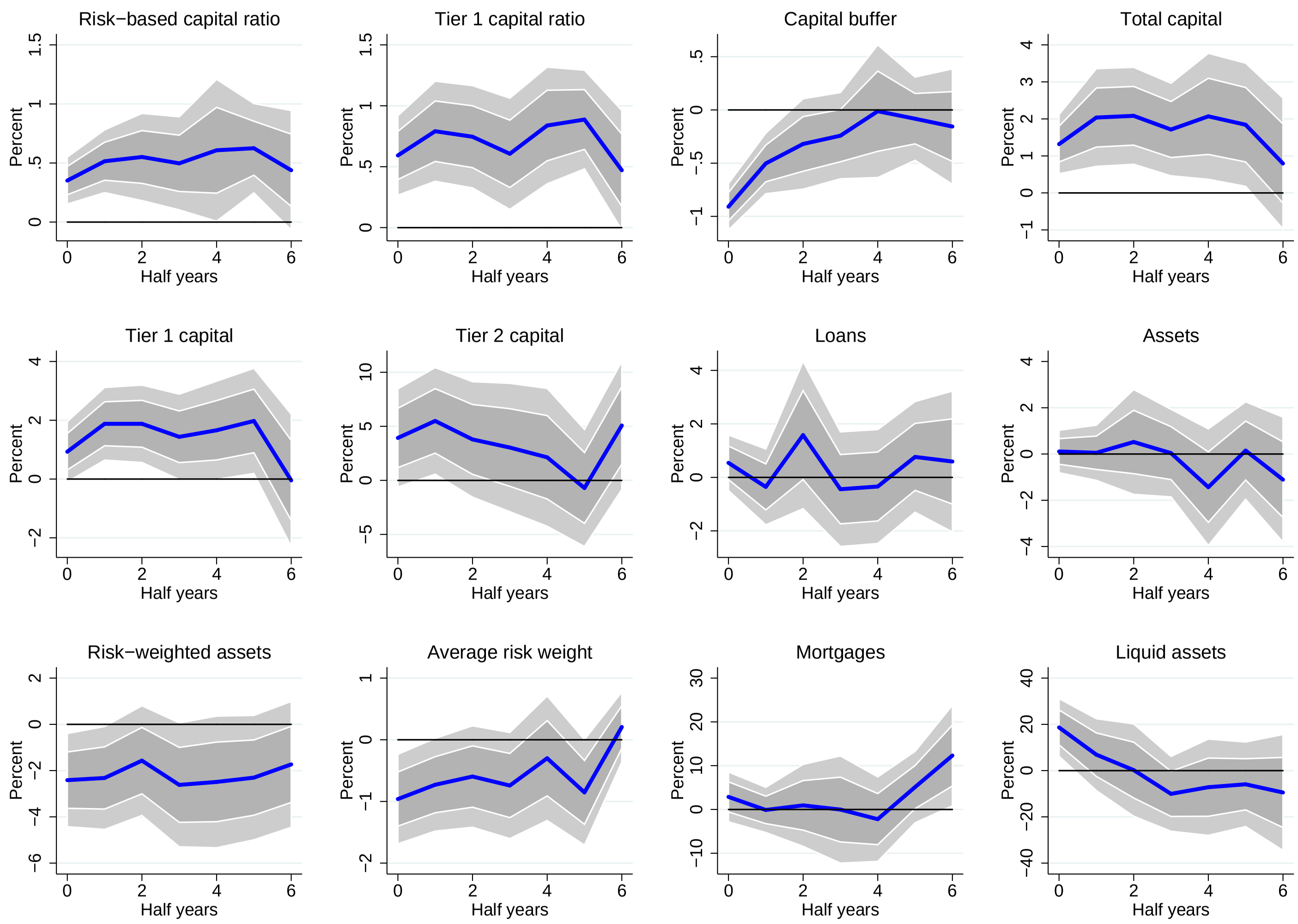
Note: this figure shows the local projection impulse responses separately based on the sign of changes in trigger ratios once bank and time fixed effects have been removed. "Decreases" (in blue) uses only observations with a negative value of this measure and the impulse response is for a 1pp *decrease* in capital requirements. "Increases" (in red) uses only positive values and the impulse response is for a 1pp *increase* in capital requirements. Impulse responses follow estimation of Equation 2.1 for twelve bank variables. Table 2.4 gives further details on the dependent variables. 68% confidence intervals are shown.

FIGURE 2.12: Impulse responses to a 1pp capital requirement increase: Small vs. large banks



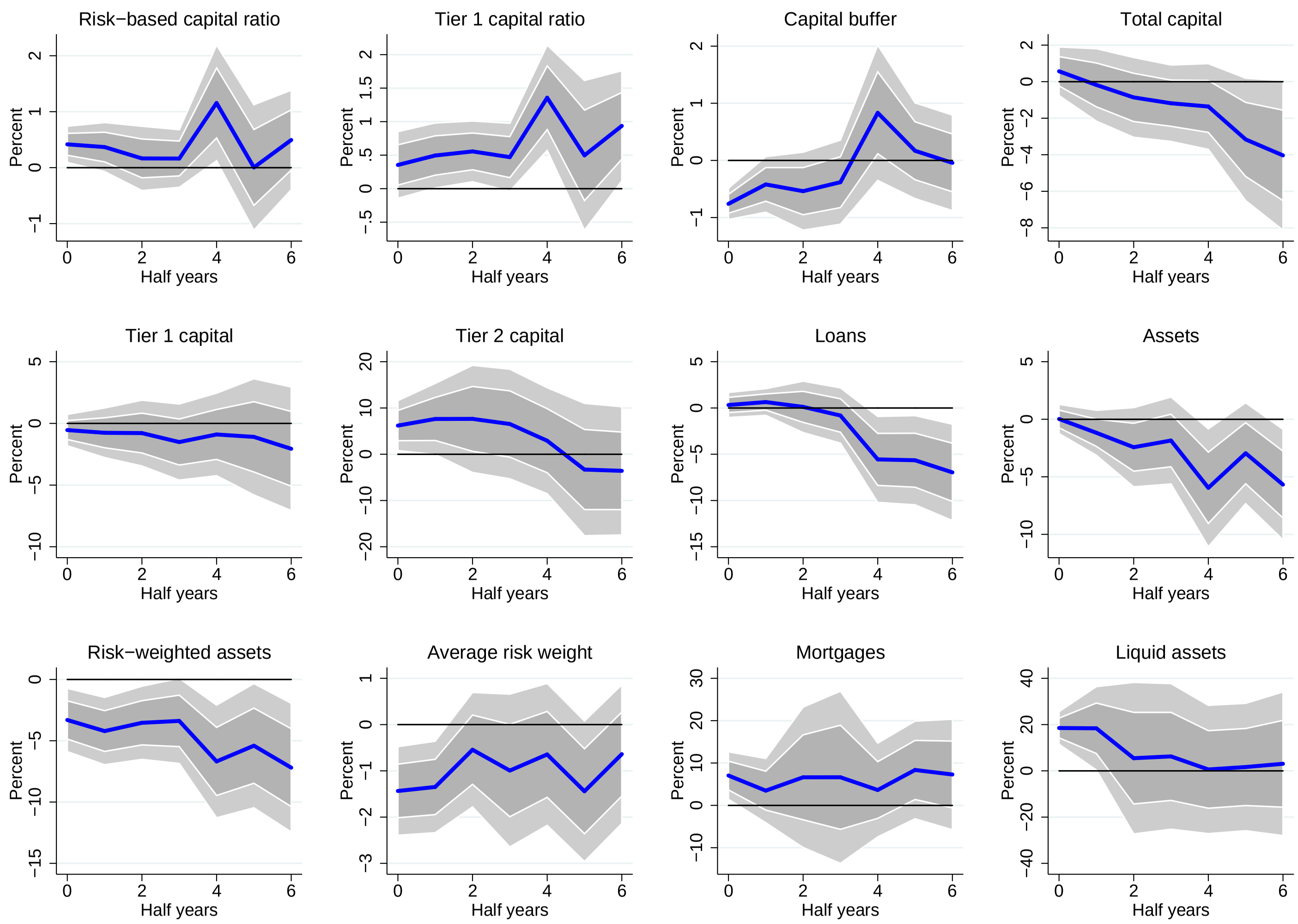
Note: this figure shows the local projection impulse responses to a 1pp capital requirement increase separately based on the size of the bank. "Small banks" (in blue) uses observations where the bank size is below the median of total assets from a given half-year period, while "large banks" (in red) covers banks with total assets above the median value. Impulse responses follow estimation of Equation 2.1 for twelve bank variables. Table 2.4 gives further details on the dependent variables. 68% confidence intervals are shown.

Figure 2.13: Impulse responses to a 1pp capital requirement increase using only observations for which a requirement change occurred



Note: this figure shows the impulse responses from estimation of Equation 2.1 for twelve bank variables following a 1pp increase in capital requirements using only those observations for which a change in trigger ratios actually took place. A change has occurred if the half-year change in trigger ratio exceeds 0.1pps in absolute value. Table 2.4 gives details on the dependent variables. 68% and 90% confidence intervals are shown.

FIGURE 2.14: Impulse responses to a 1pp capital requirement increase using only observations for which no change in requirements has occurred in the past 12 months



Note: this figure shows the impulse responses from estimation of Equation 2.1 for twelve bank variables following a 1pp increase in capital requirements using only those observations where no change in trigger ratios has occurred in the preceding 12 months. A regression of the change in capital ratios on the lasso-selected balance sheet variables (see Column 9 of Table 2.10) and bank and time fixed effects is used to strip the change in capital ratios of these variables. The residual is used in place of $\Delta trigger_{i,t}$ in Equation 2.1. A change has occurred if the half-year change in trigger ratio exceeds 0.1pps in absolute value. Table 2.4 gives details on the dependent variables. 68% and 90% confidence intervals are shown.

TABLE 2.5: Summary statistics by time period

	Pre-crisis			Post-crisis			Difference	p-value
	N	Mean	SD	N	Mean	SD		
Trigger ratio	2,630	11.564	2.785	626	11.686	2.691	0.122	0.677
Change in trigger ratio (half year, if change)	295	-0.064	0.728	311	0.149	1.482	0.214	0.003***
Tier 1 risk-based capital ratio	2,630	19.264	16.470	626	19.039	14.322	-0.225	0.895
Total risk-based capital ratio	2,630	21.351	14.710	626	22.185	14.188	0.834	0.617
Capital buffer	2,630	9.721	13.276	626	10.297	13.643	0.576	0.716
Assets growth (half-year)	2,630	3.729	9.006	626	1.072	9.685	-2.658	0.000***
Risk-weighted assets growth (half-year)	2,630	3.436	8.419	626	0.164	10.747	-3.272	0.000***
Average risk weight	2,630	58.573	19.923	626	49.878	20.803	-8.695	0.000***
Liquid asset ratio	2,209	9.691	10.563	620	14.438	10.470	4.747	0.000***
Tier 1 leverage ratio	2,630	10.805	9.574	626	9.484	8.269	-1.321	0.202
Solvency ratio	2,630	177.506	96.493	626	186.866	106.130	9.360	0.441
Loans-to-assets ratio	2,630	51.325	27.711	626	48.994	27.163	-2.331	0.447
Tier 1 capital growth (half-year)	2,630	3.335	7.970	625	2.595	9.702	-0.740	0.103
Total capital growth (half-year)	2,630	3.356	7.772	626	1.751	8.960	-1.605	0.000***
Loans growth (half-year)	2,630	3.286	10.902	626	2.293	12.034	-0.993	0.192
Deposits growth (half-year)	2,614	3.037	16.469	613	2.592	16.816	-0.445	0.594
Unsecured loans growth (half-year)	2,536	3.149	14.616	619	10.454	69.700	7.305	0.011**
Residential loans growth (half-year)	2,101	4.322	27.042	520	40.968	446.971	36.646	0.057*

Note: this table provides summary statistics separately for two time periods. The first three columns show the sample size, mean and standard deviation for the pre-crisis period, defined to be up to and including 2007H1. Columns 4-6 show the corresponding statistics for the post-crisis period, which covers 2007H2-2013H2 inclusive. Column 7 shows the difference in the two means and Column 8 gives the p-value testing equality of means. “Change in trigger ratio (half year, if change)” is the half-year change in the trigger ratio using only those observations where a capital requirement change occurred (a change is coded as having occurred if the half-year change exceeds 0.1pps in absolute value). “Average risk weight” is the ratio between risk-weighted assets and total assets. “Liquid asset ratio” is the ratio of liquid assets to total assets, where liquid assets here are defined as high quality liquid assets as well as credit to other financial institutions, debt securities and equity shares. “Solvency ratio” is the ratio between total regulatory capital and total required capital. “Unsecured loans growth” is the half-year growth of loans not secured on residential property.

Table 2.6: Summary statistics by direction of change in trigger ratio

	Decrease			Increase			Difference	p-value
	N	Mean	SD	N	Mean	SD		
Proportion of post-crisis observations	1,988	0.163	0.369	1,268	0.238	0.426	0.075	0.000***
Proportion of large banks	1,988	0.511	0.500	1,268	0.474	0.500	-0.037	0.049**
Trigger ratio	1,988	11.398	2.804	1,268	11.883	2.682	0.485	0.000***
Change in trigger ratio (half year, if change)	317	-0.729	0.788	289	0.895	0.929	1.624	0.000***
Tier 1 risk-based capital ratio	1,988	18.858	15.961	1,268	19.790	16.248	0.932	0.108
Total risk-based capital ratio	1,988	20.980	14.461	1,268	22.345	14.814	1.366	0.009***
Capital buffer	1,988	9.494	13.212	1,268	10.360	13.545	0.865	0.065*
Assets growth (half-year)	1,988	3.654	9.179	1,268	2.535	9.192	-1.119	0.000***
Risk-weighted assets growth (half-year)	1,988	3.545	8.873	1,268	1.650	9.092	-1.894	0.000***
Average risk weight	1,988	57.079	19.926	1,268	56.624	21.082	-0.455	0.554
Liquid asset ratio	1,685	10.382	10.608	1,144	11.245	10.872	0.863	0.064*
Tier 1 leverage ratio	1,988	10.325	9.201	1,268	10.906	9.574	0.581	0.079*
Solvency ratio	1,988	178.091	98.847	1,268	181.210	97.888	3.119	0.356
Loans-to-assets ratio	1,988	51.592	27.670	1,268	49.754	27.510	-1.838	0.077*
Tier 1 capital growth (half-year)	1,987	3.274	8.028	1,268	3.066	8.793	-0.208	0.537
Total capital growth (half-year)	1,988	3.062	7.816	1,268	3.025	8.377	-0.036	0.913
Loans growth (half-year)	1,988	3.255	11.130	1,268	2.846	11.139	-0.409	0.337
Deposits growth (half-year)	1,976	2.959	16.570	1,251	2.942	16.482	-0.018	0.977
Unsecured loans growth (half-year)	1,911	3.312	17.841	1,244	6.535	48.754	3.223	0.020**
Residential loans growth (half-year)	1,631	9.946	179.687	990	14.305	231.805	4.359	0.613

Note: this table provides summary statistics separately based on the sign of changes in trigger ratios once bank and time fixed effects have been removed. Columns 1-3 and 4-6 show the sample size, mean and standard deviation for decreases and increases respectively. Column 7 shows the difference in the two means and Column 8 gives the p-value testing equality of means. “Proportion of post-crisis observations” gives the proportion of observations in the relevant subgroup that are from the post-crisis period (2007H2 and beyond). “Proportion of large banks” gives the proportion of observations in the subgroup that relate to large banks, where a bank is coded as large if its total assets exceed the median value across all banks in that half year period. “Change in trigger ratio (half year, if change)” is the half-year change in the trigger ratio using only those observations where a capital requirement change occurred (i.e. if the half-year change exceeds 0.1pps in absolute value). “Average risk weight” is the ratio between risk-weighted assets and total assets. “Liquid asset ratio” is the ratio of liquid assets to total assets, where liquid assets here are defined as high quality liquid assets as well as credit to other financial institutions, debt securities and equity shares. “Solvency ratio” is the ratio between total regulatory capital and total required capital. “Unsecured loans growth” is the half-year growth of loans not secured on residential property.

Table 2.7: Summary statistics by size of bank

	Small banks			Large banks			Difference	p-value
	N	Mean	SD	N	Mean	SD		
Proportion of post-crisis observations	1,639	0.193	0.395	1,617	0.191	0.393	-0.002	0.942
Trigger ratio	1,639	12.778	2.724	1,617	10.380	2.238	-2.398	0.000***
Change in trigger ratio (half year, if change)	273	0.042	1.370	333	0.048	1.001	0.006	0.937
Tier 1 risk-based capital ratio	1,639	26.373	18.786	1,617	11.971	7.646	-14.402	0.000***
Total risk-based capital ratio	1,639	27.820	17.320	1,617	15.117	6.691	-12.703	0.000***
Capital buffer	1,639	14.865	16.377	1,617	4.730	5.938	-10.135	0.000***
Assets growth (half-year)	1,639	2.850	9.298	1,617	3.592	9.084	0.742	0.097*
Risk-weighted assets growth (half-year)	1,639	2.711	9.541	1,617	2.904	8.428	0.193	0.667
Average risk weight	1,639	57.917	21.755	1,617	55.873	18.840	-2.044	0.475
Liquid asset ratio	1,371	9.342	11.535	1,458	12.037	9.722	2.695	0.059*
Tier 1 leverage ratio	1,639	14.701	10.850	1,617	6.344	4.647	-8.358	0.000***
Solvency ratio	1,639	212.821	120.591	1,617	145.335	49.969	-67.486	0.000***
Loans-to-assets ratio	1,639	47.242	28.785	1,617	54.560	25.875	7.318	0.061*
Tier 1 capital growth (half-year)	1,638	2.466	8.229	1,617	3.929	8.377	1.463	0.000***
Total capital growth (half-year)	1,639	2.496	8.039	1,617	3.606	8.000	1.110	0.002***
Loans growth (half-year)	1,639	3.017	11.978	1,617	3.175	10.209	0.158	0.779
Deposits growth (half-year)	1,614	3.030	17.937	1,613	2.875	15.003	-0.154	0.823
Unsecured loans growth (half-year)	1,609	4.552	36.576	1,546	4.614	30.307	0.062	0.967
Residential loans growth (half-year)	1,204	19.344	294.514	1,417	5.006	30.330	-14.338	0.104

Note: this table provides summary statistics separately based on the size of the bank. A bank is "large" if its total assets exceed the median value across all banks in that half-year period. Columns 1-3 and 4-6 show the sample size, mean and standard deviation for small and large banks respectively. Column 7 shows the difference in the two means and Column 8 gives the p-value testing equality of means. "Proportion of post-crisis observations" gives the proportion of observations in the relevant subgroup that are from the post-crisis period (defined as 2007H2 and beyond). "Change in trigger ratio (half year, if change)" is the half-year change in the trigger ratio using only those observations where a capital requirement change occurred (a change is coded as having occurred if the half-year change exceeds 0.1pps in absolute value). "Average risk weight" is the ratio between risk-weighted assets and total assets. "Liquid asset ratio" is the ratio of liquid assets to total assets, where liquid assets here are defined as high quality liquid assets as well as credit to other financial institutions, debt securities and equity shares. "Solvency ratio" is the ratio between total regulatory capital and total required capital. "Unsecured loans growth" is the half-year growth of loans not secured on residential property.

Table 2.8: Lasso-selected reaction functions for pre- and post-crisis periods

	(1) Pre-crisis	(2) Post-crisis
Constant	-0.00548	-0.10178
	(0.008)	(0.142)
Time fixed effects	Yes	Yes
Bank fixed effects	Yes	Yes
Observations	2630	626
Number of banking groups	193	97
R-squared	0.10	0.05

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses. This table shows the lasso estimation following Section 6.1. The first column ("Pre-crisis") uses only observations from before the financial crisis (up to and including 2007H1), while the second column ("Post-crisis") uses from 2007H2 and beyond. The dependent variable is the half-year change in trigger ratio. The variables that appear in the lasso estimation are given in Table 2.3. An intercept and bank and time fixed effects are included.

Table 2.9: Lasso-selected reaction functions for small and large banks

	(1) Small	(2) Large
Constant	0.01661	-0.00553
	(0.035)	(0.011)
Time fixed effects	Yes	Yes
Bank fixed effects	Yes	Yes
Observations	1639	1617
Number of banking groups	131	111
R-squared	0.08	0.14

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses. This table shows the lasso estimation following Section 6.1. The first column ("Small") uses only observations where the bank size is below the median of total assets from a given half-year period, while the second column ("Large") uses banks with total assets above the median value. The dependent variable is the half-year change in trigger ratio. The variables that appear in the lasso estimation are given in Table 2.3. An intercept and bank and time fixed effects are included.

TABLE 2.10: Robustness checks: lasso-selected reaction functions under different lasso approaches and samples

	(1) Adaptive	(2) Rigorous	(3) EBIC	(4) $K = 5$	(5) $K = 8$	(6) $K = 12$	(7) $K = 15$	(8) If change	(9) No change in 12m
Change in average risk weight									-0.00472
									(0.004)
Non-core Tier 1 capital ratio change									0.00644
									(0.006)
Risk-weighted assets growth									-0.00353**
									(0.001)
Unsecured loans growth									0.00220***
									(0.001)
Non-performing loans growth									0.00002***
									(0.000)
Constant	-0.00361	-0.00361	-0.00361	-0.00361	-0.00361	-0.00361	-0.00361	-0.11748	-0.04976
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.250)	(0.033)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3256	3256	3256	3256	3256	3256	3256	606	1420
Number of banking groups	212	212	212	212	212	212	212	141	112
R-squared	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.16	0.15

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parentheses. This table shows the lasso estimation based on different lasso approaches and samples. The dependent variable is the half-year change in trigger ratio. The variables that appear in the lasso estimation are given in Table 2.3. An intercept and bank and time fixed effects are included. Columns 1, 2 and 3 show the selected model based on using adaptive lasso (Zou, 2006), clustered rigorous lasso (Belloni et al., 2016) and the extended Bayesian information criterion (Chen and Chen, 2008) respectively. Columns 4-7 give the selected model using K-fold cross validation for different K. Column 8 uses only observations for which a change in trigger ratio actually took place, where this is determined by whether the half-year change in trigger ratio exceeds 0.1pps in absolute value. Column 9 uses only observations for which no change in trigger ratio has occurred in the previous 12 months. For both Columns 8 and 9, the lasso approach follows the baseline approach of the standard lasso estimator and K-fold cross validation with $K = 10$.

Chapter 3

Do bank capital requirements affect lending? Evidence from Basel I

1 Introduction

Following the Global Financial Crisis of 2008, policymakers have been increasingly concerned with the build-up of risks on bank balance sheets, leading to a sharp rise in the number of financial policy committees and use of macroprudential policies around the world (Edge and Liang, 2019). One of the most central tools in the policymaker's toolkit is bank capital requirements. Such regulation aims to ensure that banks hold sufficient capital against their assets so that they can withstand losses, remain solvent and continue lending to the real economy should a crisis ensue. Given the important role that banks play in real economic activity, it is important to understand how banks adjust to these policies. While banks can increase their capital ratios by raising capital, which can be considered "good deleveraging" (Gropp et al., 2018), they could alternatively reduce the size of their balance sheet by cutting their loans. The latter could reduce aggregate demand and therefore have adverse macroeconomic effects (Hanson et al., 2011). This begs the question: is there a trade-off faced when deciding upon financial regulation between greater financial stability in the future and macroeconomic outcomes today? This paper sheds light on these potential costs of financial regulation by exploiting the US implementation of the Basel I reforms to study how tighter capital requirements affect bank lending.

Announced by the Basel Committee on Banking Supervision (BCBS) in July 1988, Basel I was the first major macroprudential policy of its kind. In particular, the reforms introduced the concept of risk-based capital regulation. Under Basel I, banks were required to have at least an 8% risk-based capital ratio (i.e. ratio of capital to risk-weighted assets) by the end of 1992. Risk-weighting means assets are grouped based on their perceived credit risk, meaning riskier assets such as commercial & industrial loans and consumer loans were given a higher risk weighting and thus required more capital to be held against them. Therefore, in generating a significant shift in US capital regulation, Basel I provides a useful setting for studying the impact of capital requirements.

The paper contributes to the literature in various ways: first, a natural way to identify the impact of the regulation is to exploit the heterogeneous exposures across banks to the new regulation and adopt a difference-in-differences setup. Some banks would already be conforming to the regulation prior to the announcement and thus should be less affected compared to banks with pre-reform capital ratios below 8%. Undertaking a difference-in-differences approach requires knowing risk-based capital ratios prior to Basel; however, data on risk-based capital ratios is not reported in the regulatory returns until 1990, two years after the announcement of Basel I. Consequently, many papers that have studied Basel I are pure cross-sectional studies.[1] For example, Hall (1993) studies the relationship between risk-based capital ratios and lending growth using data from 1990-91 and shows that banks that were undercapitalised at the start of the 1990s had lower lending growth. A concern of using purely cross-sectional variation is that there is no control period against which to compare this post-announcement behaviour.[2] It could be the case that undercapitalised banks generally have lower loan growth irrespective of the Basel reforms. To overcome this, I predict capital ratios going back to 1984. Using these pre-reform ratios, I divide banks into treatment and control groups. Second, the level of a bank's capital ratio prior to the Basel reform is not random. Some banks may have become undercapitalised due to a non-regulatory shock, and failure to properly account for confounding factors could invalidate the common trends assumption required for difference-in-differences analysis and therefore bias the estimated effects. Existing papers on Basel I have not examined the pre-reform behaviour of "treated" banks, making them vulnerable to such biases. I therefore provide a characterisation of undercapitalised banks to understand who they are. In particular, I show that undercapitalised banks were on average larger, had higher

[1] An exception is Berger and Udell (1994), who do include a control period. They do not document how their pre-1990 risk-based capital ratios are constructed, though they do state that a large number of assumptions are made to construct them.

[2] Other papers that face a similar concern include Johnson (1991) and Peek and Rosengren (1992, 1995).

loan-to-asset ratios and often specialised in residential and real estate loans. The latter characteristic can partly reflect the reversal of the US real estate boom that began in the mid-1980s. In line with this, many savings and loans associations (S&Ls) were undercapitalised. By identifying these characteristics and confounding events such as the S&L crisis, I choose the sample of banks appropriately to mitigate the threats of such factors. Third, existing papers do not consider the dynamics of adjustment to changes in regulation - do banks adjust fully and instantly to any new regulations or do they respond gradually? Basel I is a useful setting for studying dynamic adjustment because banks were given four years to conform to the policy. I therefore employ a dynamic difference-in-differences specification. This not only allows me to check for common pre-trends, but also enables me to study the relative behaviour of the two groups throughout the Basel transition period.

I find that undercapitalised banks increase their capital ratios gradually in response to tighter requirements rather than sharply adjusting right after the announcement or just before the final deadline. They do this by reducing the size of their balance sheet, in particular loans, with a 1 percentage point (pp) increase in capital ratios leading to a 2.5% and 5% fall in total assets and loans respectively. Digging into the types of loans, I show that reductions in lending were driven by commercial and industrial (C&I) and non-residential real estate loans with no effect on residential mortgages. This finding could be attributed to the lower risk weight of 50% given to residential mortgages compared to the 100% risk weight applied to other loan types.

The remainder of the paper is organised as follows: Section 2 provides an overview of US capital regulation. Section 3 describes the data used in the empirical analysis and Section 4 outlines the methodology used. Section 5 shows the results and Section 6 provides robustness checks, while Section 7 concludes.

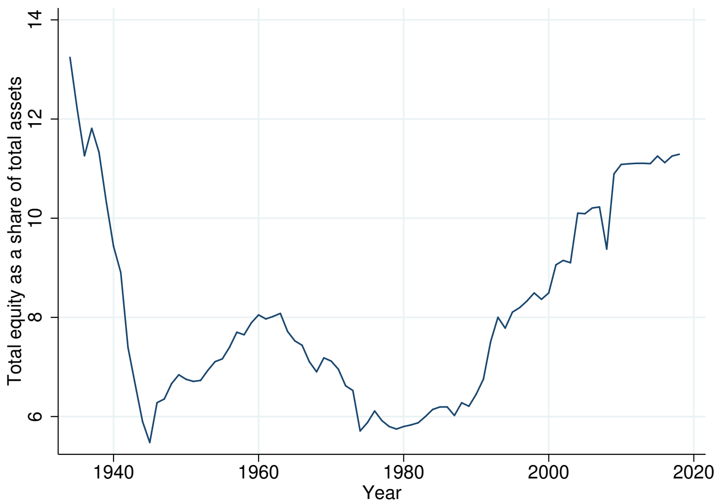
2 Institutional background

2.1 Pre-Basel I capital regulation in the US

Until the early 1980s, numeric minimum capital ratio requirements were not set in regulation. Instead, bank regulators opted for a more subjective approach tailored to individual institutions. A prevailing belief was that focusing on capital ratios would lead to a less-comprehensive analysis of a bank's loss-bearing capacity (Wall, 2014). Indeed, the 1978 FDIC Manual of Examination Policies said: *"...capital ratios...are but a first approxima-*

tion of a bank's ability to withstand adversity. A low capital ratio by itself is no more conclusive of a bank's weakness than a high ratio is of its invulnerability." (Federal Deposit Insurance Corporation, 2003). Capital ratios began to fall in the early 1980s as the oil crisis and the Volcker period of high interest rates triggered a recession. Figure 3.1 plots the equity-to-total assets ratio for the US commercial banking sector as a whole, and illustrates the drop in capital ratios starting from the early 1960s and a trough at around 1980. As a result, explicit capital requirements were introduced in 1981 in the form of a minimum leverage ratio of capital to average total assets.[3]

FIGURE 3.1: Aggregate US unweighted capital ratios



Note: this figure plots total equity as a share of total assets for the US commercial banking sector as a whole. Annual data is obtained from the FDIC Historical Bank Data.

While these steps represented a large shift in regulation from a judgement-based supervisory approach to a more explicit rule-based approach, some flaws emerged. The leverage ratios used did not adjust for the riskiness of assets on the balance sheet. In addition, the rules incentivised banks to shift more into off-balance sheet activities. Banks also complained about the unequal international playing field with US banks facing higher capital requirements relative to competing banks overseas, in particular Japanese banks (Kapstein (1989, 1991)). These drawbacks fuelled discussions at the international level and resulted in the 1988 Basel Capital Accord.

[3]Each regulatory agency had a different definition of what bank capital consists of. The Federal Reserve Board and the Office of the Comptroller of the Currency imposed a minimum primary capital leverage ratio (of adjusted total assets) of 6% for community banks and 5% for larger regional institutions. Primary capital included mainly equity and loan loss reserves. Instead, the Federal Deposit Insurance Corporation (FDIC) used a threshold capital-to-assets ratio of 6% and a minimum ratio of 5%. By 1985, all banks were required to have a primary capital ratio of 5.5% (Federal Deposit Insurance Corporation, 2003).

2.2 Basel I

Basel I required banks to achieve a minimum ratio of total regulatory capital to risk-weighted assets of 8%.[4] To compute risk-weighted assets, assets were assigned into four broad risk-weight categories based on their perceived credit risk. Very low risk assets such as cash and claims guaranteed by the US government were given a 0% risk weighting, while the 20% risk category included claims guaranteed by government-sponsored agencies, most notably Fannie Mae and Freddie Mac. Residential mortgages had a 50% risk weight. The 100% risk weight category included commercial and industrial loans, loans to individuals and non-residential real estate loans.[5] Basel I also introduced a 4% Tier 1 capital requirement (as a proportion of risk-weighted assets). In this paper, I focus on the total capital requirement given that all banks who fail the Tier 1 capital requirement would also fail the total capital requirement because of an additional rule that Tier 2 capital must not exceed Tier 1 capital.[6] However, the reverse need not be true (i.e. banks that fail to achieve an 8% total risk-based capital ratio could still have a 4% Tier 1 capital ratio).[7]

Banks were not required to adjust to the new regulation instantly. While Basel I was announced by the Basel Committee in July 1988, banks did not need to fully conform until end-1992. I therefore focus on the transition period from 1988Q3 to 1992Q4. This timeline is summarised in Table 3.1. A further impetus for banks to increase their capital ratios was given by the Prompt Corrective Action provision as directed by the Federal Deposit Insurance Corporation Improvement Act of 1991. This classified institutions into five categories based on their risk-based capital ratios, and gave supervisors certain powers (e.g., suspending dividends) if banks fell into the low capitalisation categories. Banks needed an 8% risk-based capital ratio to be "adequately capitalised" (the second highest category), which kept them safe from supervisory action.[8] It is difficult to separate the impact of the Basel regulation from the tougher supervisory approach; however, one can argue that the two go hand-in-hand. For capital regulation to be effective, it needs to be

[4]Total capital is the sum of Tier 1 and Tier 2 capital. Tier 1 capital included common equity, some preferred stock and minority interest in consolidated subsidiaries minus goodwill. Tier 2 capital included loan-loss reserves, subordinated debt and other preferred and convertible stock.

[5]Off-balance-sheet items were also given a risk weight. For further details on the definitions and asset allocations, see Basel Committee on Banking Supervision (1988) and Department of the Treasury (1989).

[6]There are some further restrictions on Tier 2 capital, e.g., loan-loss reserves were limited to 1.25% of risk-weighted assets. See Department of the Treasury (1989) for further details.

[7]There was also a Tier 1 leverage requirement which varied based on the bank's CAMEL rating, a regulatory rating from 1-5 given to banks by their supervisors. Banks with a CAMEL rating of 1 (the highest rating) had to obtain a ratio of Tier 1 capital to total unweighted assets of 3%, and banks with lower ratings had to maintain a ratio of 1-2 percentage points higher. As CAMEL ratings are confidential, I cannot study this requirement explicitly in the paper.

[8]For further details on these provisions, see Jones and King (1995).

enforced. The Prompt Corrective Action provision provided this enforcement mechanism.

TABLE 3.1: Timeline of Basel I

July 1988	Basel Committee publish the Basel Accord
August 1988	Governors of Federal Reserve System approve Basel Accord
January 1989	US regulators publish final rules for implementation
31^{st} December 1990	Intermediate 7.25% risk-based capital ratio requirement
31^{st} December 1992	Full implementation of 8% risk-based capital ratio requirement

3 Data

I use bank-level data from the Statistics on Depository Institutions (SDI) compiled and made publicly available by the Federal Deposit Insurance Corporation (FDIC). This dataset contains quarterly financial data, as well as structural information on the bank. The SDI dataset is particularly beneficial for my setting for a variety of reasons: first, it covers all FDIC-insured financial institutions, and so my analysis is not restricted to large and/or listed banks. Second, the data begins in 1984Q1, thus providing a suitably long control period. Third, a central difficulty in studying Basel I is that risk-based capital ratios are not reported until 1990Q1, reflecting the fact that this variable was introduced by the Basel regulation. The SDI contains a large number of financial variables that I will use to construct a measure of risk-based capital ratios going back to 1984Q1. Fourth, the dataset contains a breakdown of loan portfolios by type (e.g., consumer, residential, real estate and commercial & industrial), which will allow me to study the effect of capital requirements on different loan types.

4 Method

My empirical approach takes two main steps: first, I obtain an estimate of bank capital ratios going back to 1984Q1 by using post-Basel data (1993Q1-1999Q4) to construct a mapping between a large number of financial variables and risk-based capital ratios. Second, I employ a dynamic difference-in-differences framework to study whether undercapitalised banks, identified from the pre-reform capital ratios estimated in the first step, adjusted their balance sheet differently to capitalised banks following the Basel reforms.

4.1 Constructing risk-based capital ratios for the 1980s

The objective of the first step is to understand the capitalisation of a bank before Basel such that a treatment and control group for the difference-in-differences analysis can be constructed. As noted in Section 3, risk-based capital ratios are not reported until 1990Q1. Furthermore, the degree of aggregation of balance sheet variables in the 1980s is such that it is difficult to manually construct risk-based capital ratios without making strong assumptions. Although data on total equity and unweighted assets are available, using leverage ratios (i.e. the ratio of total equity to unweighted assets) may not be appropriate ex-ante to identify which banks are undercapitalised with respect to the Basel I regulations.[9] To see this, consider two banks A and B with the same level of both equity and assets. If bank A holds riskier assets such as consumer loans, while bank B holds safe assets like cash and US Treasuries, then bank A is more undercapitalised in risk-weighted terms. As a result, comparing banks based on unweighted capital ratios would not be appropriate.

To obtain a measure of risk-based capital ratios going back to the 1980s, I estimate the following parametric model using data from 1993Q1-1999Q4, which is the period after the full implementation of Basel I:

$$\text{Capital ratio}_{it} = \frac{\beta' X_{it}}{\gamma' X_{it}} + \epsilon_{it} \tag{3.1}$$

where Capital ratio$_{it}$ is the risk-based capital ratio of bank i at time t and X_{it} is a set of 47 financial variables.[10] The procedure for selecting the set of financial variables is described in Section B.1 and the full list of variables in X_{it} is given in Table 3.3. One can think of this process as trying to back out the implicit risk weights on each financial variable in X_{it}. Using the estimated coefficients, I take the risk-based capital ratio of bank i at time t as:

$$\widehat{\text{Capital ratio}}_{it} = \frac{\hat{\beta}' X_{it}}{\hat{\gamma}' X_{it}}$$

[9] Peek and Rosengren (1992, 1995) use leverage ratios rather than risk-based ratios in their analysis.

[10] To reduce the impact of anomalous values for capital ratios (e.g., due to reporting errors), I only consider banks with ratios between -10% and 100%, which covers 98% of all ratio observations in this period. Note that it is possible for risk-based capital ratios to be negative, e.g., because of deductions made as part of the regulatory calculations or undivided profits are negative. However, negative ratios are rare with only 0.1% of observations falling into this category.

4.2 Difference-in-differences estimation

To study the impact of the Basel I regulation, I use a dynamic difference-in-differences framework, which on first glance has notable differences compared to standard applications. A typical setup generally involves a group that are not treated at all (the control group) and a group that is given some treatment (the treatment group). Applying this framework to Basel I is somewhat unusual because all banks are in some way likely to be affected by the new capital regulation - there is no completely untreated set of banks. Undercapitalised banks are affected because they are forced to increase their capital ratios to conform to the new regulation. However, capitalised banks can also be affected; for example, they may also decide to build a larger capital buffer to avoid the risk of falling below the 8% threshold, or there could be be general equilibrium effects. This thus begs the question of how to interpret the resulting estimates. We should interpret the difference-in-differences estimates as the *relative* impact of a binding capital constraint on undercapitalised banks against capitalised banks. It will tell us the impact of being forced to increase your capital ratio to the Basel threshold relative to banks that do not have to. As such, it captures not only the direct adjustment from undercapitalised banks, but also any indirect effects on capitalised banks such as general equilibrium effects.

While using difference-in-differences restricts attention to relative effects, there are advantages to using this framework for studying the impact of capital requirements: first, a difficulty in studying the impact on capital ratios/requirements on lending is that loans are a large component of total assets and thus enter into the denominator of capital ratios. As such, a simple contemporaneous regression of loans on capital ratios would lead to biased estimates due to this reverse causality, i.e. increases in lending directly reduce capital ratios. One solution to this is to identify an instrumental variable; however, these can be difficult to find. Using a difference-in-differences framework avoids the need to find a valid instrument for capital ratios as I define control and treatment groups based on pre-Basel capital ratios and use treatment status as my variable of interest rather than the time-varying ratios themselves. This approach is used by Juelsrud and Getz Wold (2020), who study the implementation of Basel III capital requirements in Norway. Second, the use of a dynamic setup gives insights into the speed of adjustment, which is particularly interesting in this setting because banks were given four years to conform to the new regulation.

I allocate banks with risk-based capital ratios below 8% in 1988Q1 to the treatment group, while banks with higher ratios are placed in the control group. The baseline specification is as follows:

$$y_{ist} = \alpha_i + \delta_{st} + \sum_{k=1984Q1}^{1992Q4} \beta_k \mathbb{1}(t = k) \cdot T_i + \epsilon_{ist} \quad (3.2)$$

where y_{ist} is the outcome of interest for bank i located in state s at time t.[11] α_i denote bank fixed effects, while T_i is a binary variable equal to 1 if bank i belongs to the treated group and zero otherwise. The quarter 1988Q4, which is the year before the Basel transition period begins, is omitted, and so the coefficients of interest β_k give the quarter-specific marginal effect of being undercapitalised relative to 1988Q4 on the outcome variable y_{ist}.[12] δ_{st} are state-by-time fixed effects, which help to account for both aggregate shocks and geographical differences in economic conditions. Indeed, the sample period coincides with the early 1990s recession and various regional shocks (e.g., the oil price collapse in the mid-1980s hitting oil-producing states such as Texas).

The standard assumption in difference-in-differences estimation is common trends, whereby the post-Basel path of the outcome y_{ist} for treated and undercapitalised banks in the absence of Basel is the same as that of the capitalised banks. To provide some confidence in this assumption, I exploit the dynamic nature of the specification and test whether $\beta_k = 0$ for $k \leq 1988Q4$. These coefficients give the effect of treatment status on the outcome of interest in the periods *before* the Basel transition period. Values of zero imply common pre-trends between the two groups.

The main hypotheses to test are as follows:

Hypothesis 1: For total assets and loans and $k \geq 1989Q1$, $\beta_k < 0$ as undercapitalised banks shrink the size of their balance sheets to increase their capital ratios.

Hypothesis 2: β_k is more negative for loans with higher risk weights, namely consumer, non-residential real estate and consumer loans, which each have a 100% risk weight). These loans should fall by more than residential mortgages, which has a 50% risk weight.

[11] A full list of the dependent variables considered with a description of each is given in Table 3.4. Other than for the capital ratio, which is in percentage points, all other dependent variables are in logs when estimating Equation 3.2.

[12] I take the treatment period to be 1989Q1-1992Q4 as the US regulators published the rules on Basel I implementation in January 1989 (see Table 3.1), and so 1989Q1 is the first quarter after this publication.

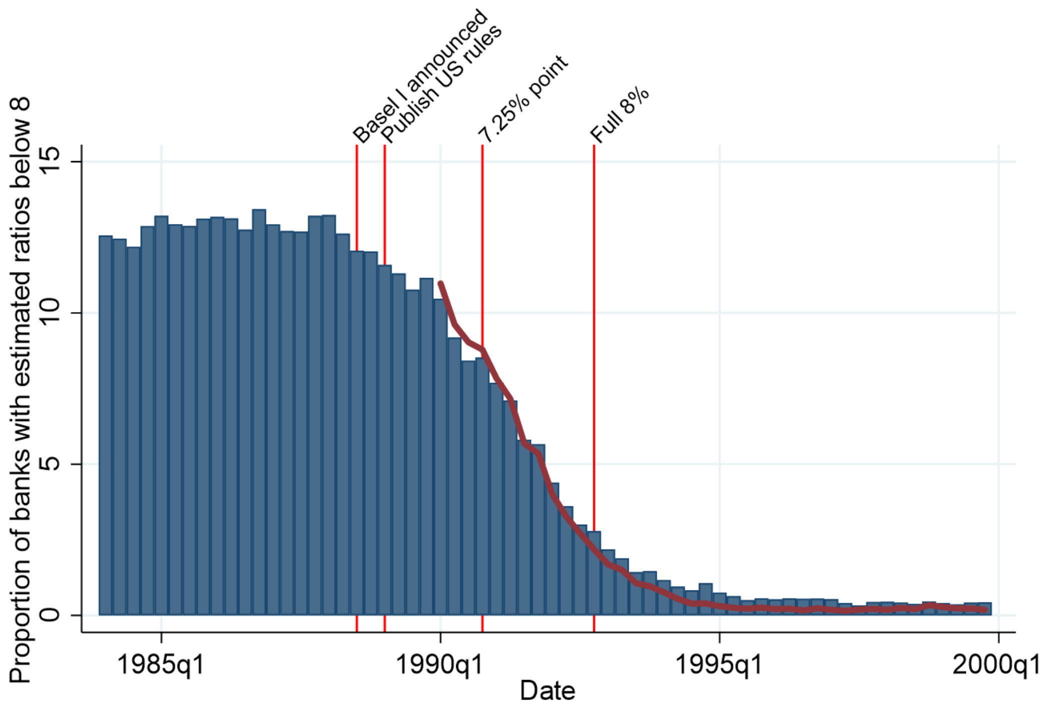
5 Results

5.1 Fit of capitalisation measure

I begin by running Equation 3.1 to obtain a mapping between financial variables and risk-based capital ratios, which I then use to obtain measures of risk-based capital ratios in the 1980s. Figure 3.11 plots the number of banks for which capital ratios can be estimated, and shows a high coverage with over 98% of banks having estimated capital ratios during the earlier period of my sample. To evaluate the fit of these capitalisation proxies, I compare these estimated capital ratios against actual risk-based capital ratios using data from 1990Q1-1992Q4. This test data was not used in estimation of Equation 3.1. A strong fit of the proxy during this period would be reassuring because 1990Q1-1992Q4 is a dynamic period for capital ratios given the new capital regulation and the early 1990s recession. Figure 3.2 shows the proportion of banks in the sample with estimated capital ratios below 8% in each quarter going back to 1984Q1. The figure also plots the actual proportion of banks with ratios below 8% as the maroon line (data only available from 1990Q1). From this figure, we can see that the period from 1990Q1-1992Q4 is indeed a dynamic one with the proportion of banks with ratios below the Basel minimum of 8% falling sharply. The estimated proportions line up well with the actual proportions. The model was estimated on data from 1993Q1-1999Q4, which Figure 3.2 shows is a relatively stable period at least in terms of the lower tail of the capital distribution. Being able to match movements during 1990Q1-1992Q4 adds confidence in the model used to estimate capital ratios going back to 1984Q1.

To more directly assess the performance of the individual estimates, Figure 3.3 shows a bin scatter plot of actual against estimated risk-based capital ratios again using test data from 1990Q1-1992Q4. The points line up almost perfectly against the 45-degree line. Figure 3.4 shows a quantile-quantile plot of actual against estimated capital ratios using the test data. The plot lines up almost perfectly with the 45-degree line. Both figures thus suggest that the model fits actual risk-based capital ratios well. To evaluate the success in correctly assigning banks into the treatment and control groups, I check the proportion of observations that would be correctly sorted in the 1990Q1-1992Q4 test data. 93.5% of observations were assigned to the correct group.

FIGURE 3.2: Proportion of banks with capital ratios below 8%



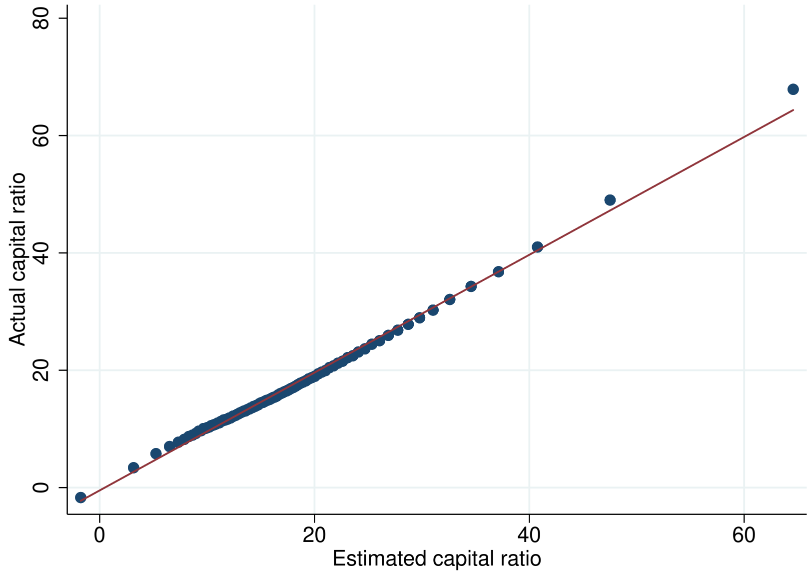
Note: this figure plots the proportion of banks with capital ratios below 8%. The bars show the proportion using the estimated capital ratios obtained by running Equation 3.1. The maroon line shows the actual proportion computed using true capital ratios (available only from 1990Q1). The red vertical lines reference key milestones of Basel I as outlined in Table 3.1.

5.2 Analysis of capitalisation in the 1980s

In this section, I provide a characterisation of undercapitalised banks during the pre-Basel period. Such analysis is important because capitalisation status prior to the reforms is not necessarily random. Indeed, undercapitalised banks may not have always been undercapitalised, e.g., they may have faced an adverse shock shortly before the new regulations were announced. Such shocks could confound the difference-in-differences analysis. In particular, they could lead to a violation of the common trends assumption. Therefore, it is important to understand the nature of the treated and control groups to be able to mitigate the threat from any confounding factors.

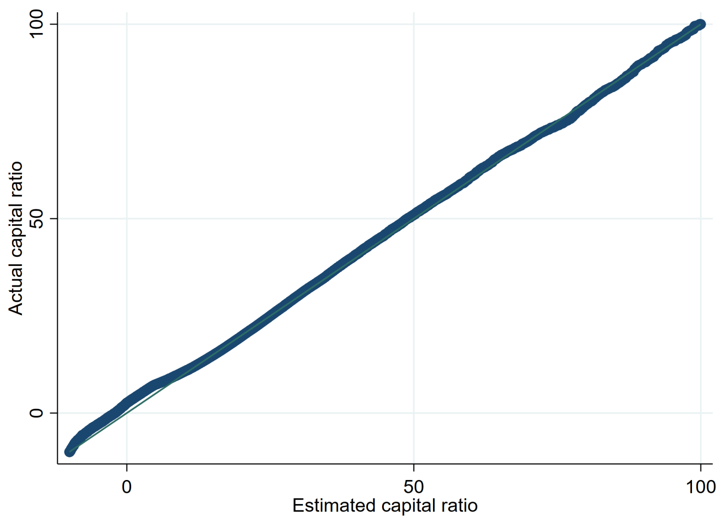
Figure 3.2 shows that the incidence of undercapitalisation was rather steady in the lead-up to Basel I. Around 13% of banks were undercapitalised relative to the Basel I standards when the regulations were announced in 1988, and this proportion remained steady going back to 1984. In both the actual and estimated data, we see that the decline in the proportion of undercapitalised banks was gradual. There was not a sudden drop as soon as the regulations were announced or just before the full implementation deadline. Figure

Figure 3.3: Bin scatter plot of actual against estimated capital ratios using test data (1990Q1-1992Q4)



Note: this figure shows a bin scatter plot of actual risk-based capital ratios (y-axis) against estimated risk-based capital ratios (x-axis) obtained by estimating Equation 3.1. The red diagonal line is a 45-degree line. Data used is from 1990Q1 to 1992Q4. 100 quantiles are used in producing the bin scatter plot. The sample is restricted to banks with estimated ratios between -10% and 100%, which covers 98.4% of banks for which estimated capital ratios could be computed.

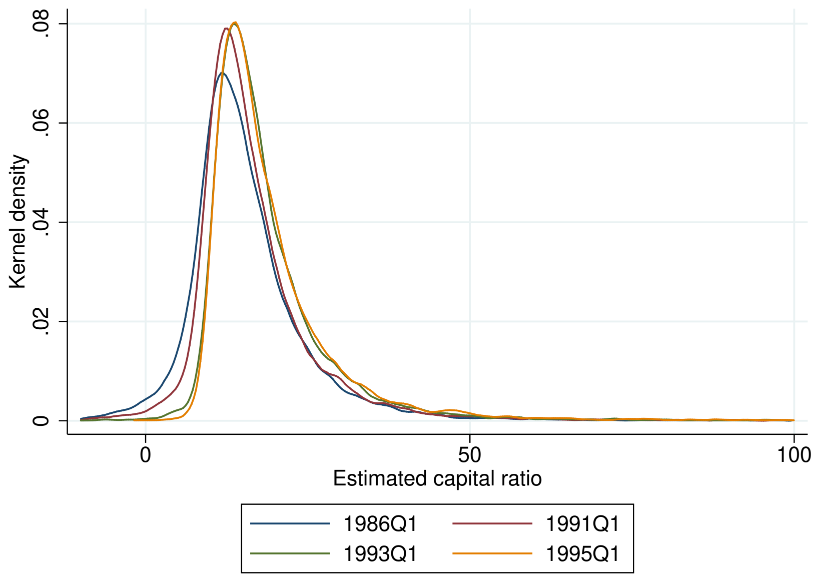
Figure 3.4: Quantile-quantile plot of actual against estimated capital ratios using test data (1990Q1-1992Q4)



Note: this figure shows a quantile-quantile plot of actual risk-based capital ratios (y-axis) against estimated risk-based capital ratios (x-axis) obtained by estimating Equation 3.1. The green diagonal line is a 45-degree line. Data used is from 1990Q1 to 1992Q4. The sample is restricted to banks with estimated ratios between -10% and 100%, which covers 98.4% of banks for which estimated capital ratios could be computed.

3.5 shows a kernel density plot of the estimated capital ratios at four points: 1986Q1 (pre-Basel), 1991Q1 (after intermediate 7.25% point), 1993Q1 (first quarter after full Basel implementation) and 1995Q1 (two years after full implementation).[13] There is a significant rightward shift in the distribution of capital ratios from 1986Q1 to 1993Q1, and this shift is particularly strong on the left tail of the distribution. The distribution remains stable from 1993Q1 to 1995Q1, and so it can be argued that much of the shift in capital ratios can be linked to the Basel regulation.

FIGURE 3.5: Kernel density plot of estimated capital ratios



Note: this figure shows a kernel density plot of estimated capital ratios obtained from running Equation 3.1. The blue, red, green and yellow lines correspond to 1986Q1, 1991Q1, 1993Q1 and 1995Q1 respectively. The sample is restricted to banks with estimated ratios between -10% and 100%. The Epanechnikov kernel function is used to produce the kernel densities.

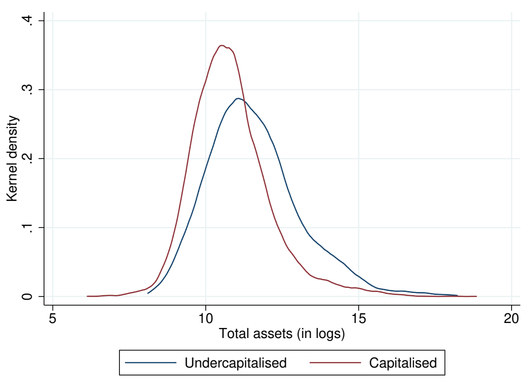
I now identify some characteristics of undercapitalised banks relative to capitalised banks. Figure 3.6a shows a kernel density plot of total assets for undercapitalised and capitalised banks separately. The figure shows that undercapitalised banks were generally larger. The same pattern emerges for total loans (see Figure 3.12). Digging into the type of loan, an interesting feature is that many undercapitalised banks seemed to specialise in real estate loans, in particular residential mortgages. Figure 3.6b shows a kernel density plot of the residential loans share (as a proportion of total loans). For undercapitalised banks, there is a clear mass of banks with very high residential loans shares.[14] I subsequently check whether savings and loans associations (S&Ls) dominate the undercapitalised group. To do this, I divide banks into 20 quantiles based on their 1988Q1 estimated capital ratios. Figure

[13]The figure is almost identical if actual rather than estimated capital ratios are used for 1991Q1, 1993Q1 and 1995Q1, which again gives confidence in the constructed capital ratio measure.

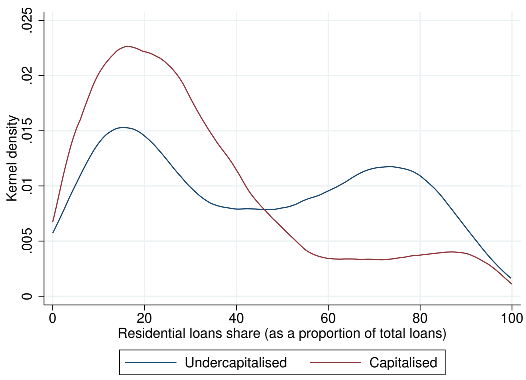
[14]A similar pattern also holds for real estate loans as a whole as shown in Figure 3.13.

3.14 shows the proportion of banks within each quantile group that are S&Ls. Just under 70% and 50% of banks in the first and second quantiles respectively were S&Ls compared to a population average of 15.5%. The undercapitalisation of S&Ls can be linked to a variety of causes (see Ely, 2008); however, a key feature is that S&Ls faced a maturity mismatch between assets (primarily fixed-rate mortgages) and liabilities (customer deposits). As such, the Volcker period of high interest rates drove many S&Ls into insolvency. Congress allowed some deregulation in the early 1980s in the hope that this would fuel growth of S&Ls and help them get out of insolvency. Examples include allowing S&Ls to hold more non-residential loans and adjustable rate mortgages and removing maximum LTV limits. These looser regulations allowed S&Ls to delve into riskier lending; however, the reversal of the real estate from the mid-1980s reignited the insolvency concerns and helped to fuel the S&L crisis. Due to the unique pre-Basel history of S&Ls, I omit them in the baseline difference-in-differences estimation.

FIGURE 3.6: Total assets and residential loans share in 1988Q1 by capitalisation status



(A) Total assets (in logs)

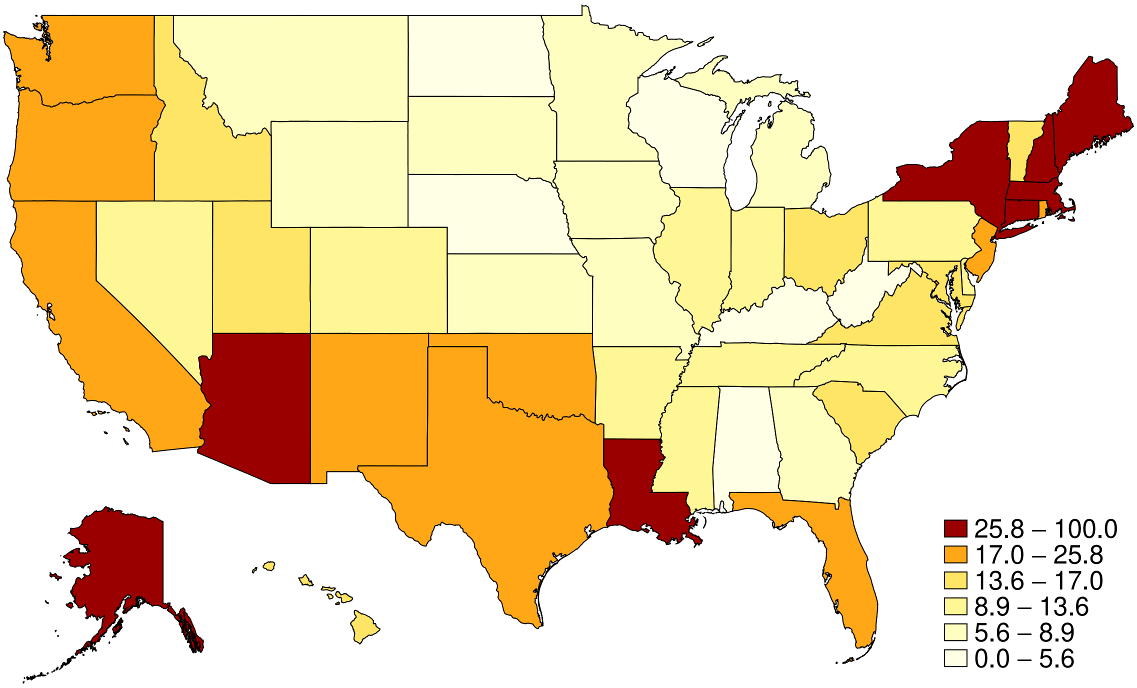


(B) Residential loan share

Note: this figure shows kernel density plots of total assets (left panel) and residential loans share as a proportion of total loans (right panel) for undercapitalised and capitalised banks. Both plots use data from 1988Q1. Undercapitalised banks are those with estimated capital ratios below 8% in 1988Q1, while capitalised banks have an estimated ratio above 8%. The estimated capital ratios are obtained from running Equation 3.1. The Epanechnikov kernel function is used to produce the kernel densities.

Figure 3.7 illustrates the proportion of banks in each state that are undercapitalised in 1988Q1 and shows that there are regional differences in the prevalence of undercapitalisation. We can see that undercapitalisation is particularly prevalent in the Southwest and Northeast. The Southwest region was reliant on oil, and so their banks faced losses on energy loans following the drop in oil prices in the mid-1980s. Furthermore, bankers tried to mitigate their losses on energy loans by investing in real estate, particularly commercial real estate, which further hit southwestern banks when real estate markets fell (Federal Deposit Insurance Corporation, 1997). The Northwest region was also hit by the decline of real estate, particularly from around 1989, and so the high prevalence of undercapitalised banks in the Northwest in 1988Q1 could be indicative of problems relating to real estate exposure that were beginning to unravel. The heterogeneity across states in capitalisation and the links to known shocks emphasises the importance of including state-by-time fixed effects in the difference-in-difference specification.

FIGURE 3.7: Proportion of banks in each state that are undercapitalised in 1988Q1



Note: this figure illustrates the proportion of all banks in a given state that are undercapitalised in 1988Q1. Undercapitalised banks are those with estimated capital ratios below 8% in 1988Q1. The colours used correspond to bins based on these proportions (ranges of each bin are given in the bottom right of the figure). Darker colours represent a higher proportion of banks being undercapitalised.

5.3 Difference-in-differences estimation

This section studies the impact of Basel I on undercapitalised banks relative to capitalised banks using the difference-in-differences framework described in Section 4.2. As the regulation applies at the bank holding company (BHC) level, I aggregate the individual banks up to this level though refer to these groups as "banks" for simplicity.[15] There are various threats to identification to consider: first, the number of banks in the US gradually declined from the mid-1980s as depicted in Figure 3.11. One may therefore be concerned about non-random exit, in particular if undercapitalised banks were more likely to exit. I therefore consider only a balanced panel in which banks need to appear in each quarter from 1984Q1 to 1995Q4. Second, Section 5.2 shows that high exposure to real estate loans characterised many undercapitalised banks, many of whom were S&Ls.

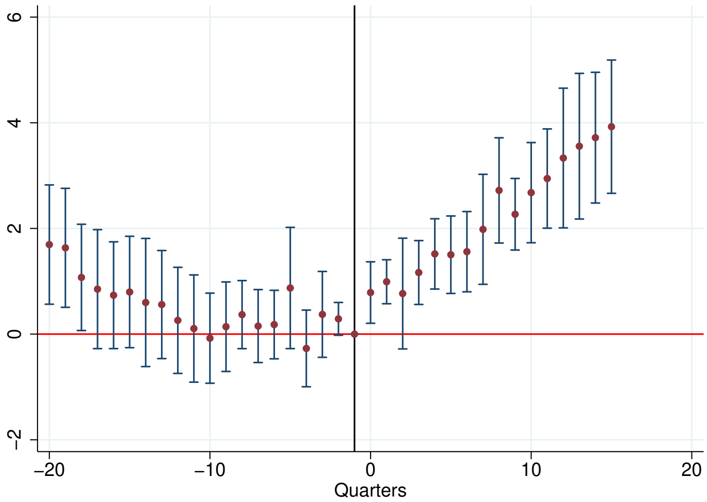
[15]To produce estimated capital ratios at the BHC level, I take the sum of the X_{it} variables for banks belonging to the same BHC and apply the coefficients obtained from estimating Equation 3.2.

This could be a threat to identification if the real estate market was not stable over the sample period. This was in fact the case with a reversal of the housing boom from the mid-1980s driving some banks to become undercapitalised. Furthermore, S&Ls exhibited very different behaviours relative to other banks in the build-up to Basel I as a result of the new regulations and lending powers granted by Congress described in Section 5.2. I thus drop BHCs with at least one S&L from the baseline specification. As a robustness check, in Section 6, I go further and also exclude all banks with a real estate loan share (as a proportion of total loans) above 50% before the introduction of Basel I. Third, while some banks may have been persistently undercapitalised in the years before Basel I, some banks may have faced an (idiosyncratic) shock that caused them to be undercapitalised as of 1988Q1. Banks could similarly experience positive shocks. I therefore restrict attention to those banks that experience an absolute change in their capital ratios no larger than 2 percentage points from 1984Q1 to 1987Q4. While this does significantly constrain the sample size, this helps to ensure that the remaining banks in the sample are free from large idiosyncratic shocks that could confound estimates. There are 1645 BHCs in the resulting sample, of which 31 are undercapitalised. Table 3.2 gives summary statistics for the two groups.

Figure 3.8 plots the β_k estimates with capital ratios as the dependent variable. The coefficients are generally insignificant in the quarters leading up to the start of the Basel transition period. The positive and significant coefficients during the Basel transition period suggest that undercapitalised banks increased their capital ratios by more than capitalised banks. The coefficients rise over time, which indicates that undercapitalised banks increased their capital ratios gradually over the transition period rather than instantly adjusting following the announcement or only increasing their ratios close to the final deadline. By 1992Q4, undercapitalised banks had increased their ratios by 4pps more relative to capitalised banks.

I now test the first hypothesis from Section 4.2, namely whether undercapitalised banks shrink the size of their balance sheet as part of their adjustment to tighter capital regulation. Figure 3.9 plots the coefficient estimates for total assets and loans. In both cases, the estimates are insignificantly different from zero in most quarters during the pre-Basel period. As with capital ratios, reductions in total assets and loans were gradual over the transition period. By the end of 1992, undercapitalised banks appear to have reduced the size of their balance sheet by about 10% relative to capitalised banks. Given that undercapitalised banks increased their capital ratios by about 4pps more over this period, a simple back-of-the-envelope calculation suggests that a 1pp increase in capital

FIGURE 3.8: Event study estimates for capital ratios

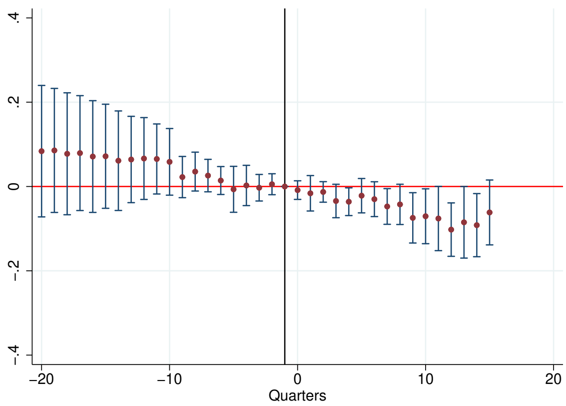


Note: this figure plots the β_k estimates from estimation of Equation 3.2 where the dependent variable y_{ist} is the risk-based capital ratio. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. Capital ratios are estimated following Equation 3.1, and are winsorised at the 1^{st} and 99^{th} percentiles of each quarter prior to estimation. Undercapitalised banks are those with a risk-based capital ratio below 8% in 1988Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

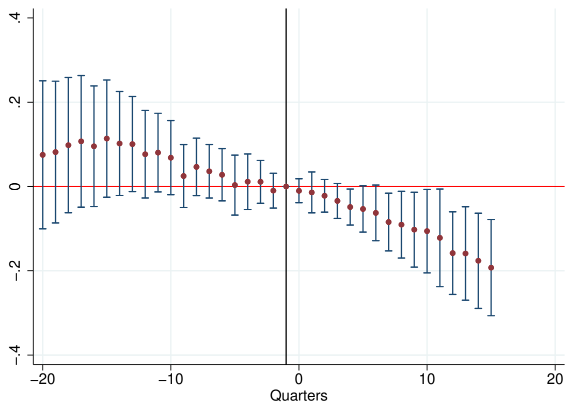
ratios translates to a $\frac{10}{4} = 2.5\%$ fall in total assets. The coefficient estimates are larger in magnitude for total loans and reach -0.2 by 1992Q4. An equivalent calculation therefore suggests that a 1pp increase in capital ratios causes a 5% fall in total loans.

FIGURE 3.9: Event study estimates for total assets and total loans

(A) Total assets



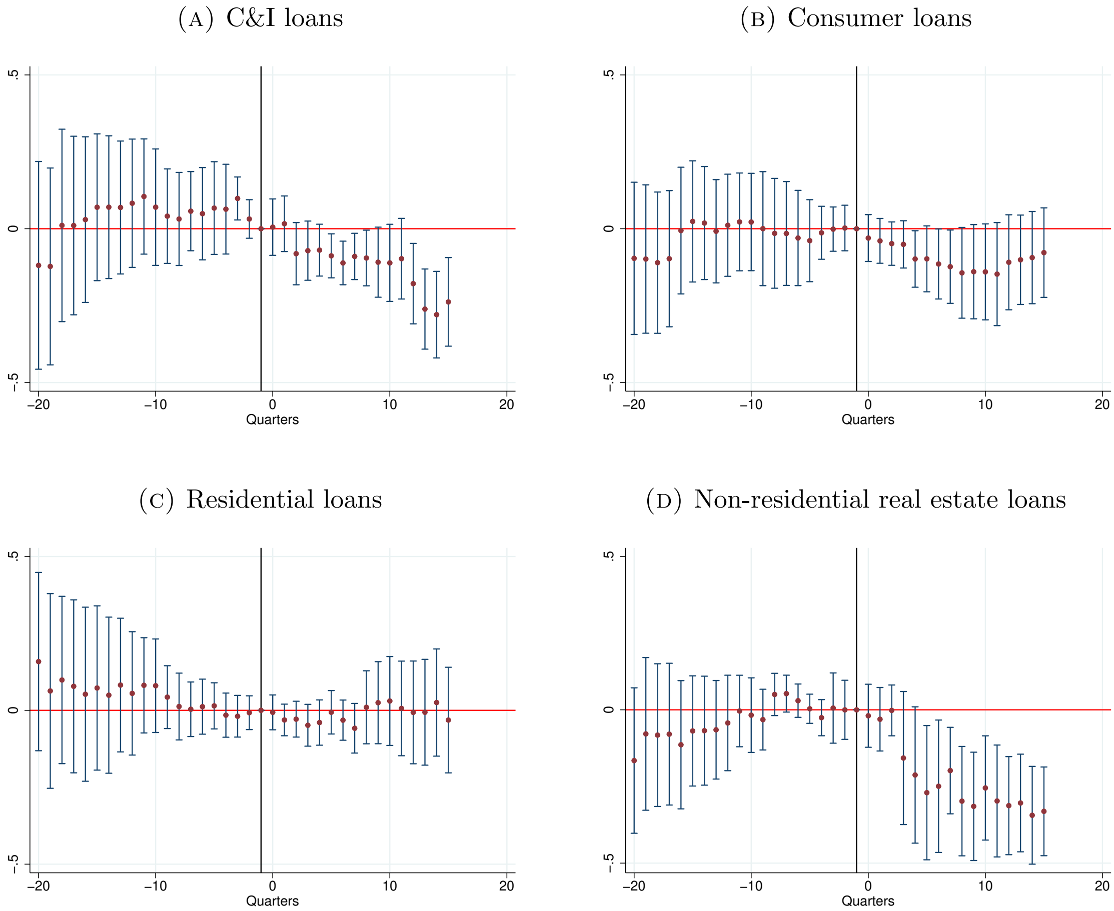
(B) Total loans



Note: this figure plots the β_k estimates from estimation of Equation 3.2. The left panel uses total assets (in logs) as the dependent variable, while the right panel uses total loans (in logs). Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q1. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

The next step is to understand whether specific loan types are affected more than others. As noted in the second hypothesis, residential mortgages received a lower risk weight (50%) compared to other loan types (100%), and so one may expect banks to cut resi-

Figure 3.10: Event study estimates by individual loan types



Note: this figure plots the β_k estimates from estimation of Equation 3.2 where the dependent variable y_{ist} is an individual loan type (commercial & industrial, consumer, residential and non-residential real estate loans). Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q1. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

dential mortgages by less. Figure 3.10 gives the difference-in-differences estimates for four individual loan types, namely commercial & industrial (C&I), consumer, non-residential real estate and residential real estate loans. In virtually all cases, the coefficient estimates for $k < 1988Q4$ are insignificant. The strongest effects appear to be in C&I and non-residential real estate loans with coefficients reaching -0.24 and -0.33 in 1992Q4 respectively. These imply that a 1pp increase in capital ratios is associated with 6% and 8.3% falls in C&I and non-residential real estate loans respectively. In accordance with the second hypothesis, there are no significant effects on residential real estate loans.

6 Robustness

I run a variety of robustness checks to verify these findings: first, I check that the results are not sensitive to the quarter in which treatment status is defined. I thus redefine treatment status based on 1988Q2 estimated capital ratios. As discussed in Section 2.2, the Basel Committee announced Basel I in July 1988, and so the regulatory data for end-quarter 1988Q2 corresponds to 1 month before this announcement. Figure 3.15 shows that the coefficient estimates are very similar to the baseline estimation. I also redefine treatment status using 1987Q4 estimated capital ratios in Figure 3.16 and the same conclusions are reached, though significance for C&I loans is less clear.

Second, Section 5.2 notes that undercapitalised banks were typically larger, allocated a greater share of assets to lending and had high real estate exposure. One might be concerned that these characteristics may interact with time-varying economic conditions and thus confound the difference-in-differences estimates. For example, if the early 1990s recession hit larger banks more than smaller banks, then one cannot be sure that the decline in lending experienced by undercapitalised banks during the Basel transition period was due to the regulation rather than the recession. To help mitigate these concerns, I interact the values of total assets, loan-to-asset share and residential loan share in 1984Q1 with the quarter fixed effects. These variables allow for differential time trends by size and loan exposure. Figure 3.17 shows that the results are robust to inclusion of these interactions.

Third, one may be concerned that real estate is more exposed to an economic downturn than other assets, and so if undercapitalised banks tend to specialise in real estate loans, the cuts in lending could reflect conditions in the real estate market during the 1990s recession hitting undercapitalised banks more than capitalised banks rather than the impact of Basel. Whilst excluding non-S&Ls from the main analysis goes some way to dealing with overexposure issues, I re-estimate the model by also dropping banks that have a real estate loan share (as a proportion of total loans) above 50% in 1988Q1. The results are robust to this, though the effects on C&I and non-residential real estate loans are quantitatively larger relative to the baseline case (see Figure 3.18). Last, Figure 3.19 shows that the results are robust to omitting very small banks.

7 Conclusion

The announcement of the Basel I reforms in 1988 represented a major shift in capital regulation in the US by introducing a risk-based minimum requirement for capital ratios, thus providing an ideal setting for identifying the impact of tighter capital requirements. Given that risk-based capital ratios are not reported until 1990, I first obtain estimates of capital ratios going back to 1984. Using these pre-reform ratios, I show that banks that were undercapitalised with respect to the new Basel standards just prior to its announcement were on average larger and specialised in residential and real estate loans. In particular, many savings and loan associations (S&Ls) were undercapitalised. Exploiting the heterogeneous exposure to the Basel reforms across banks based on pre-reform capital ratios, I estimate a dynamic difference-in-differences model to compare the behaviour of undercapitalised and capitalised banks following the Basel announcement. I find that bank capital ratios increase gradually in response to the tightening of requirements. A 1pp regulation-induced increase in capital ratios leads to 2.5% and 5% falls in total assets and loans respectively. This reduction in lending is predominantly driven by commercial & industrial and non-residential real estate loans which fall by 6% and 8.3% following a 1pp increase in ratios respectively. Instead, there is no significant effect on residential mortgage lending. This finding is in line with the lower risk weight given to residential mortgages relative to other loan types.

There are various avenues for further research on the impacts of capital regulation: first, this paper looks at on-balance-sheet activities. However, banks may have shifted more into off-balance-sheet activities such as securitisation of loans, and this could influence the amount of loans that are actually reduced due to tighter capital regulation. Exploring such securitisation activity could be an interesting dimension of study. Second, the nature of Basel I meant that the riskiness of the *individual* borrower is irrelevant for the Basel I calculations. As a result, undercapitalised banks may increase lending to riskier borrowers, which could increase the financial vulnerability of the bank. Loan-level data linked to characteristics of the borrower is needed to assess this possibility. Third, the state of the business cycle could determine the most cost-effective way to increase capital ratios. For example, it may be less costly to issue new stock during a boom and reduce lending in a recession. The implementation of Basel I coincided with a nationwide recession, which could explain the strong effects on lending found in this paper. It would be interesting to study how balance sheet adjustments differ depending on the business cycle.

References

Basel Committee on Banking Supervision (1988): "International convergence of capital measurement and capital standards," Discussion paper, Bank of International Settlements.

Berger, A. N., and G. F. Udell (1994): "Did Risk-Based Capital Allocate Bank Credit and Cause a "Credit Crunch" in the United States?," *Journal of Money, Credit and Banking*, 26(3), 585–628.

Department of the Treasury (1989): "Risk-Based Capital Guidelines; Final Rule," *Federal Register*, 54(17), 4167–4184.

Edge, R., and J. N. Liang (2019): "New Financial Stability Governance Structures and Central Banks," Finance and Economics Discussion Series 2019-019, Board of Governors of the Federal Reserve System (U.S.).

Ely, B. (2008): "Savings and Loan Crisis," *Concise Encyclopedia of Economics (2nd ed.). Indianapolis: Library of Economics and Liberty.*

Federal Deposit Insurance Corporation (1997): *History of the Eighties–lessons for the Future: An examination of the banking crises of the 1980s and early 1990s*, vol. 1. Federal Deposit Insurance Corporation.

——— (2003): "Basel and the Evolution of Capital Regulation: Moving Forward, Looking Back," Discussion paper, Federal Deposit Insurance Corporation.

Gropp, R., T. Mosk, S. Ongena, and C. Wix (2018): "Banks Response to Higher Capital Requirements: Evidence from a Quasi-Natural Experiment," *The Review of Financial Studies*, 32, 266–299.

Hall, B. J. (1993): "How Has the Basle Accord Affected Bank Portfolios?," *Journal of the Japanese and International Economies*, 7(4), 408 – 440.

Hanson, S. G., A. K. Kashyap, and J. C. Stein (2011): "A Macroprudential Approach to Financial Regulation," *Journal of Economic Perspectives*, 25(1), 3–28.

Johnson, R. (1991): "The bank credit "crumble"," *Quarterly Review*, 16, 40–51.

Jones, D. S., and K. K. King (1995): "The implementation of prompt corrective action: An assessment," *Journal of Banking & Finance*, 19(3), 491 – 510.

Juelsrud, R. E., and E. Getz Wold (2020): "Risk-weighted capital requirements and portfolio rebalancing," *Journal of Financial Intermediation*, 41, 100806.

Kapstein, E. B. (1989): "Resolving the regulator's dilemma: international coordination of banking regulations," *International Organization*, 43(2), 323–347.

——— (1991): "Supervising International Banks: Origins and Implications of Basle Accord," Discussion paper, International Economics Section, Department of Economics, Princeton University.

Peek, J., and E. Rosengren (1995): "The Capital Crunch: Neither a Borrower nor a Lender Be," *Journal of Money, Credit and Banking*, 27(3), 625–638.

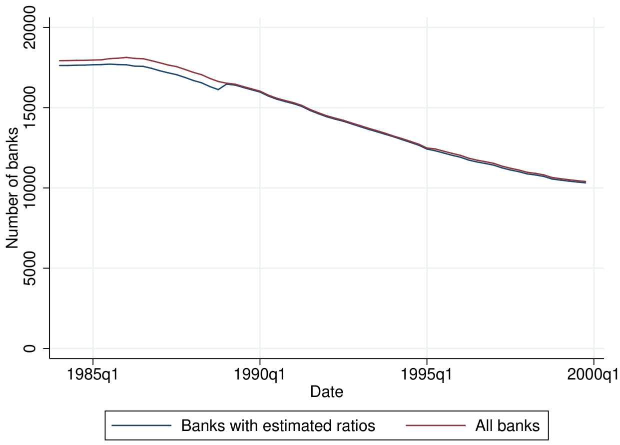
Peek, J., and E. S. Rosengren (1992): "The role of real estate in the New England credit crunch," Working Papers 92-4, Federal Reserve Bank of Boston.

Wall, L. D. (2014): "Simple Concept, Complex Regulation," Discussion paper, Federal Reserve Bank of Atlanta.

Appendix

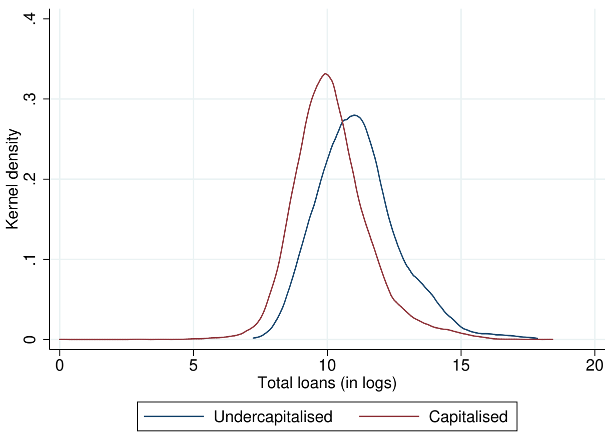
A Tables and figures

FIGURE 3.11: Number of banks



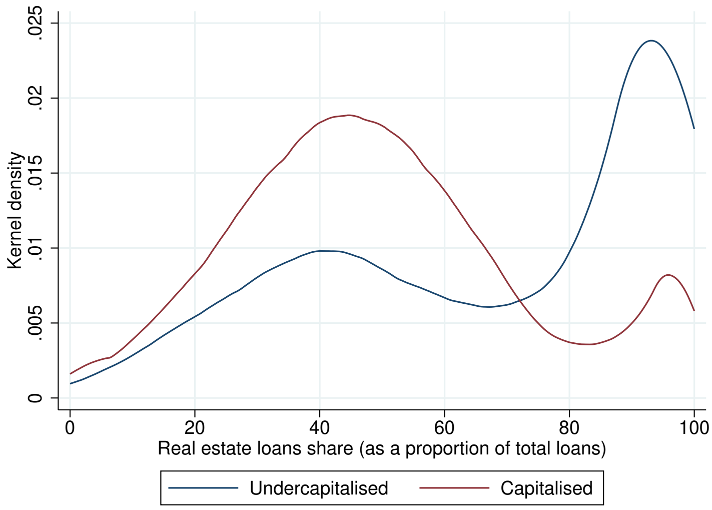
Note: this figure plots the total number of banks in the Statistics on Depository Institutions (SDI) dataset in red and the number of banks for which estimated capital ratios (using Equation 3.1) can be computed.

FIGURE 3.12: Kernel density plot of total loans by capitalisation status in 1988Q1



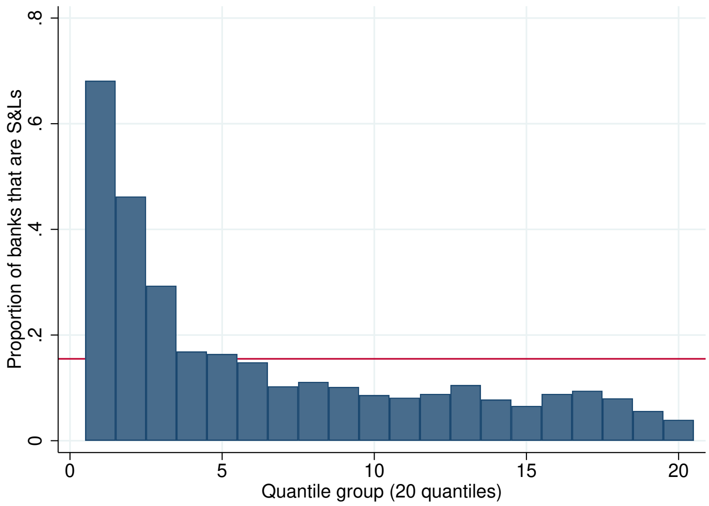
Note: this figure shows a kernel density plot of total loans (in logs) for undercapitalised and capitalised banks. Undercapitalised banks are those with estimated capital ratios below 8% in 1988Q1, while capitalised banks have an estimated ratio above 8%. The estimated capital ratios are obtained from running Equation 3.1. The Epanechnikov kernel function is used to produce the kernel densities.

Figure 3.13: Kernel density plot of real estate loans share by capitalisation status in 1988Q1



Note: this figure shows a kernel density plot of real estate loans share (as a proportion of total loans) for undercapitalised and capitalised banks. Undercapitalised banks are those with estimated capital ratios below 8% in 1988Q1, while capitalised banks have an estimated ratio above 8%. The estimated capital ratios are obtained from running Equation 3.1. The Epanechnikov kernel function is used to produce the kernel densities.

Figure 3.14: Proportion of banks that are S&Ls by capital ratio quantile in 1988Q1



Note: this figure shows the proportion of banks that are savings and loans associations (S&Ls) in each of 20 quantiles computed using 1988Q1 estimated capital ratios. The 1st quantile corresponds to the least capitalised banks in 1988Q1 (0-5th percentile), while the 20th covers the most capitalised banks (95-100th percentile). The horizontal red line corresponds to the proportion of institutions that are S&Ls across all quantiles. The estimated capital ratios are obtained from running Equation 3.1.

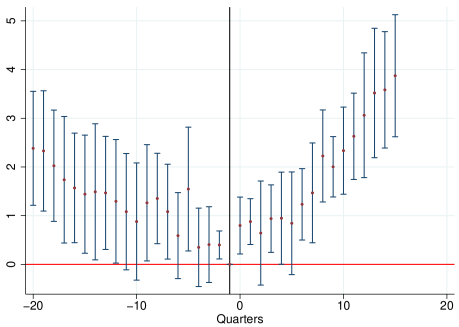
TABLE 3.2: Summary statistics by capitalisation group

	Means		
	Overcapitalised	Undercapitalised	p
Risk-based capital ratio	15.99	7.40	0.00
Equity-to-assets ratio	8.49	4.85	0.00
Total assets (in logs)	10.68	12.66	0.00
Total loans and leases (in logs)	10.01	12.26	0.00
Loans-to-asset ratio	54.03	69.23	0.00
Real estate loans share (of total loans)	35.49	41.27	0.10
Residential loans share (of total loans)	21.27	27.58	0.04
C&I loans share (of total loans)	21.64	25.18	0.15
N	1614	31	

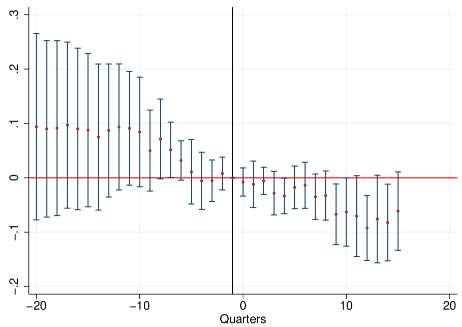
Note: this table shows mean values for a variety of variables separately for the 'Capitalised' and 'Undercapitalised' groups used for the baseline difference-in-differences specification. Means are computed using data from 1984Q1. The first column shows the group means for 'Capitalised' banks, which are those with estimated capital ratios above 8% in 1988Q1. The second column shows the 1988Q1 group means for 'Undercapitalised' banks, which have ratios below 8%. The estimated capital ratios are obtained from running Equation 3.1. The third column shows p-values of equality of means.

Figure 3.15: Event study estimates using 1988Q2 capital ratios to define capitalisation status

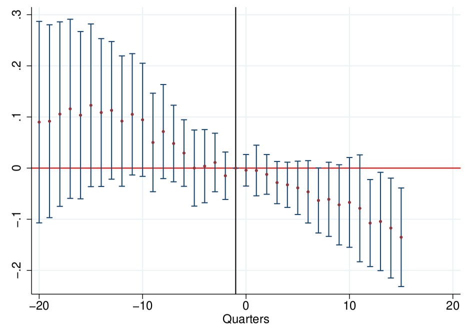
(A) Capital ratio



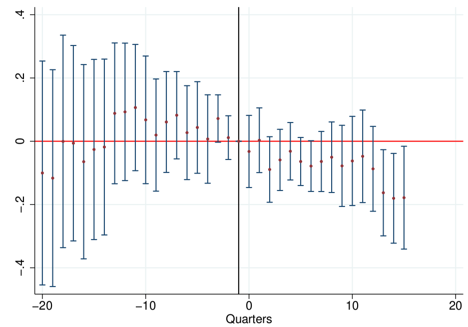
(B) Total assets



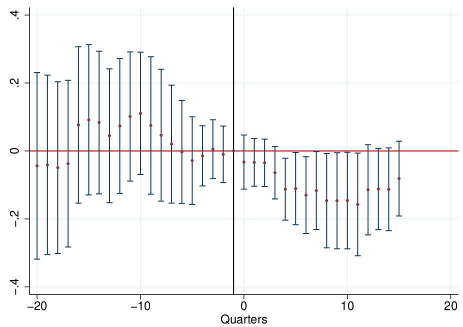
(C) Total loans



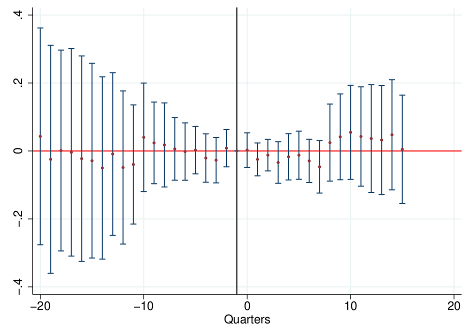
(D) C&I loans



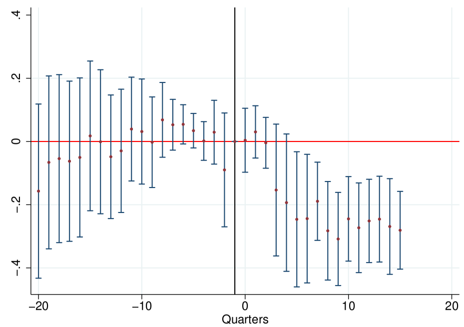
(E) Consumer loans



(F) Residential loans



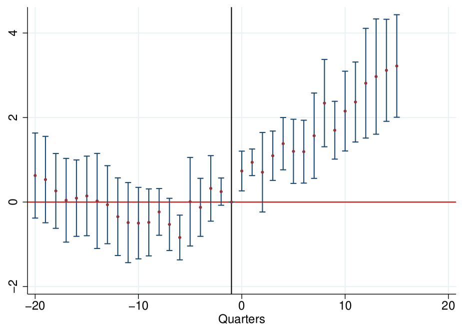
(G) Non-residential real estate loans



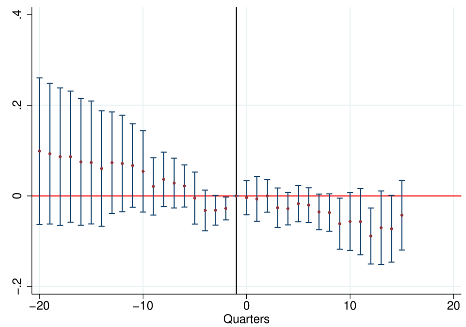
Note: this figure plots the β_k estimates from estimation of Equation 3.2 for a range of dependent variables. Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q2. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

FIGURE 3.16: Event study estimates using 1987Q4 capital ratios to define capitalisation status

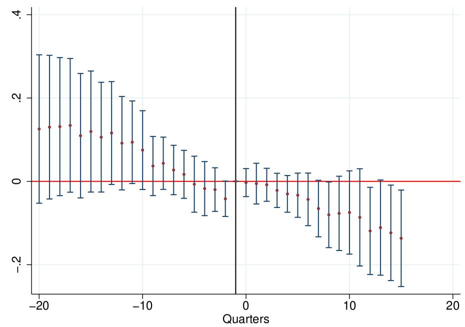
(A) Capital ratio



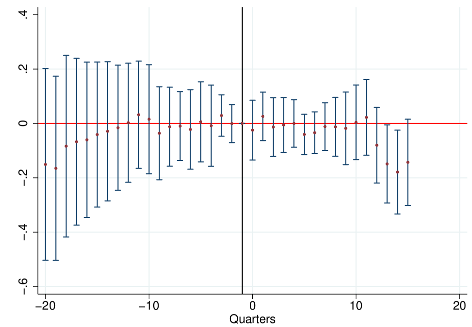
(B) Total assets



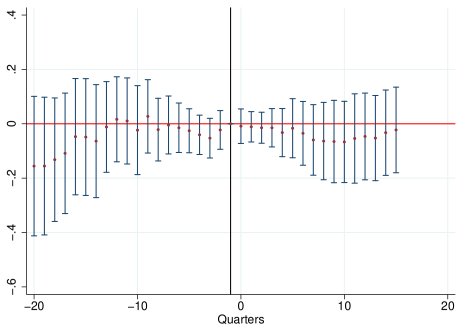
(C) Total loans



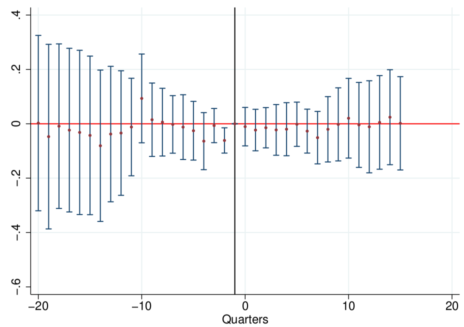
(D) C&I loans



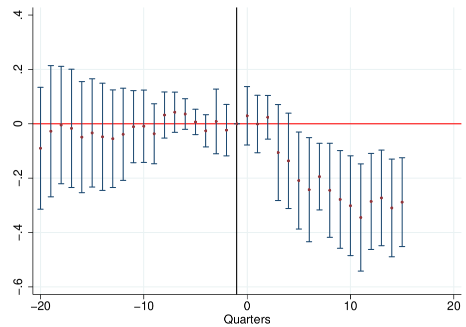
(E) Consumer loans



(F) Residential loans



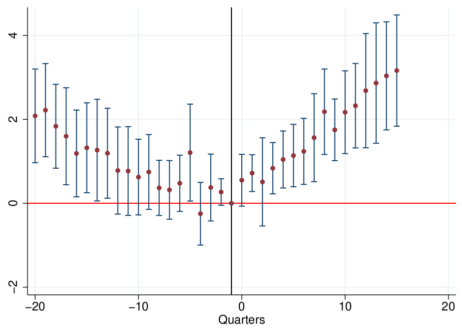
(G) Non-residential real estate loans



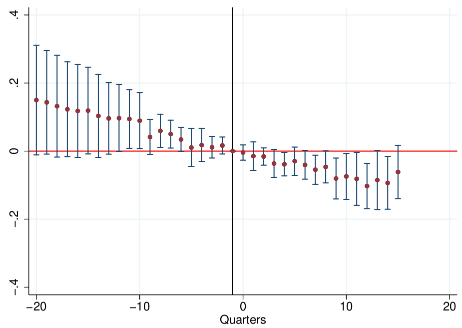
Note: this figure plots the β_k estimates from estimation of Equation 3.2 for a range of dependent variables. Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1987Q4. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

FIGURE 3.17: Event study estimates allowing for differential time trends by size, loan-to-asset share and residential loan share

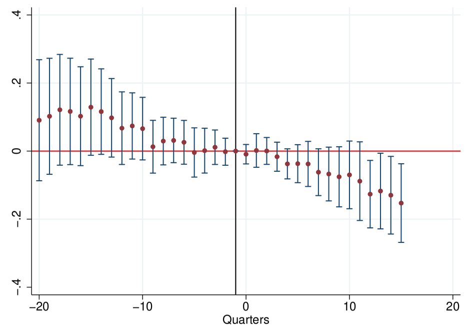
(A) Capital ratio



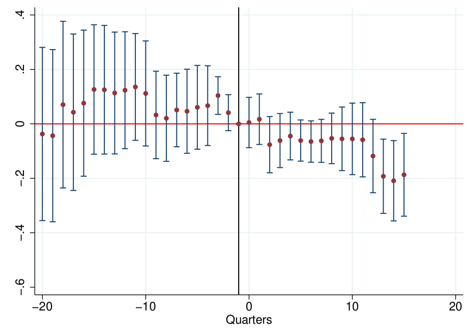
(B) Total assets



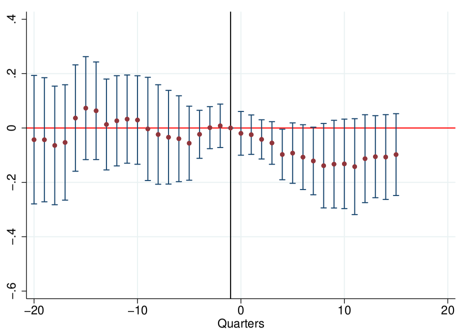
(C) Total loans



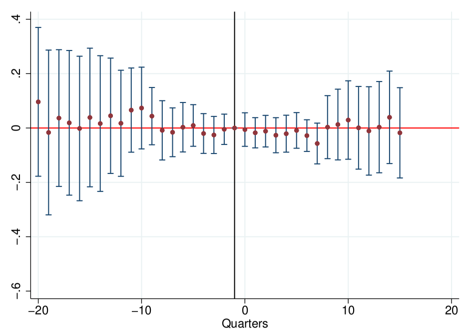
(D) C&I loans



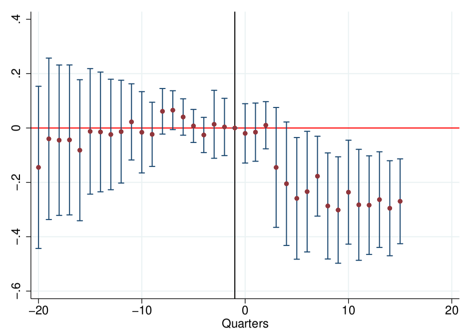
(E) Consumer loans



(F) Residential loans

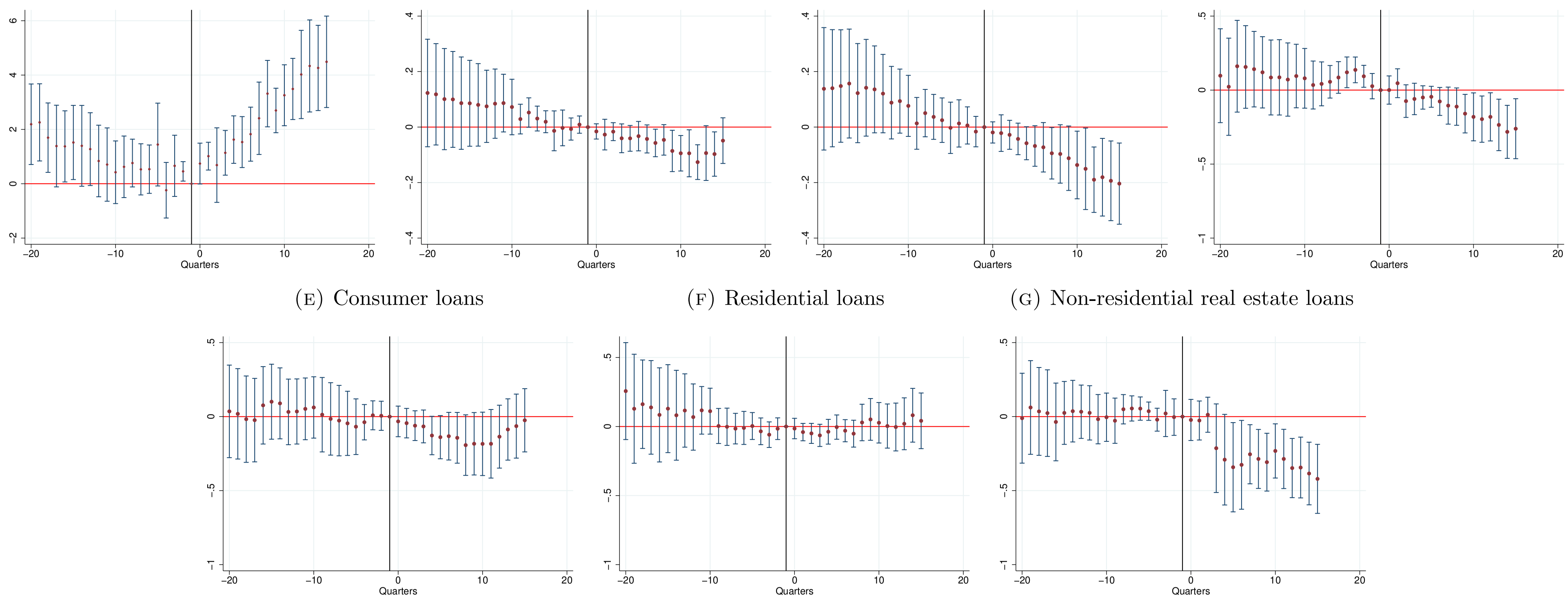


(G) Non-residential real estate loans



Note: this figure plots the β_k estimates from estimation of Equation 3.2 for a range of dependent variables, but including interactions of total assets (in logs), loan-to-asset share and residential loans share (all measured as of 1984Q1) with the quarter fixed effects. Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q1. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

FIGURE 3.18: Event study estimates omitting banks with high real estate exposure



(A) Capital ratio

(B) Total assets

(C) Total loans

(D) C&I loans

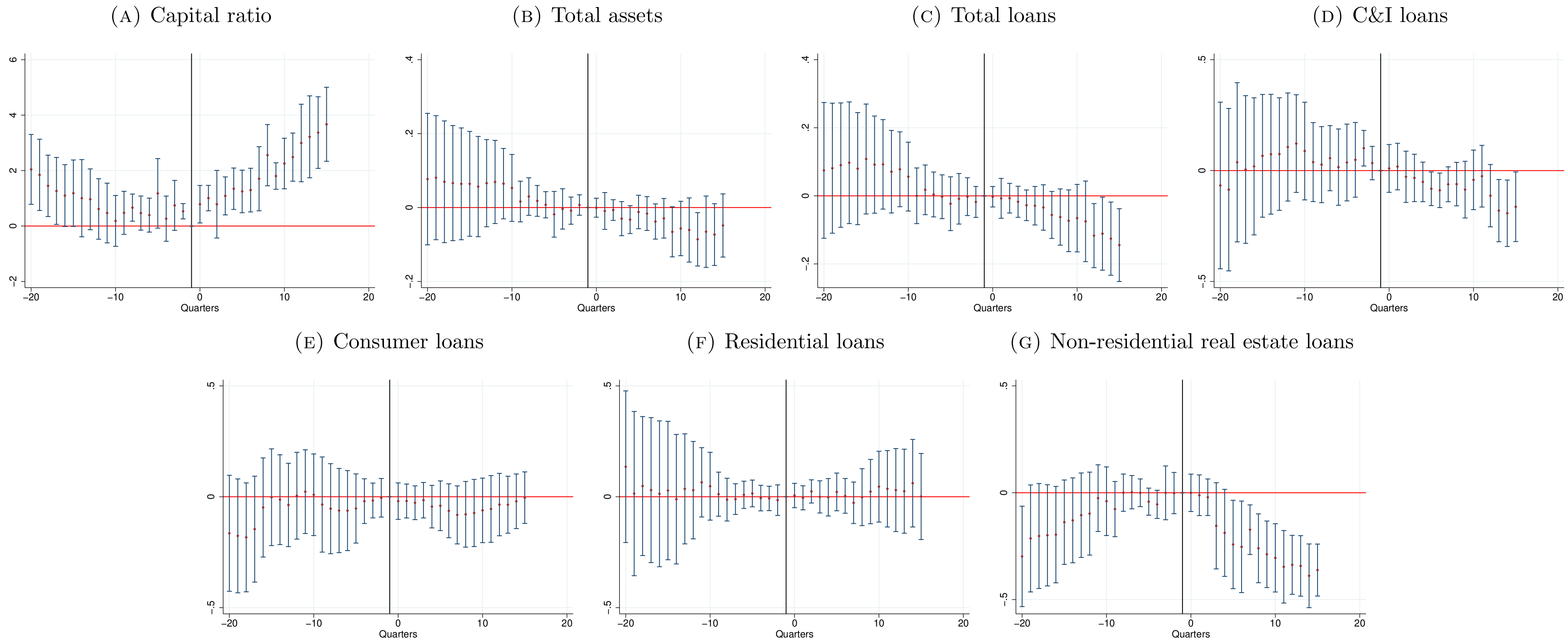
(E) Consumer loans

(F) Residential loans

(G) Non-residential real estate loans

Note: this figure plots the β_k estimates from estimation of Equation 3.2 for a range of dependent variables. Banks with real estate loan shares (relative to total loans) above 50% are excluded. Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q1. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

Figure 3.19: Event study estimates omitting very small banks



(a) Capital ratio (b) Total assets (c) Total loans (d) C&I loans

(e) Consumer loans (f) Residential loans (g) Non-residential real estate loans

Note: this figure plots the β_k estimates from estimation of Equation 3.2 for a range of dependent variables. Banks with total assets below \$25m in 1984Q1 are omitted. Undercapitalised banks are those with a risk-based capital ratio (estimated using Equation 3.1) below 8% in 1988Q1. The x-axis gives the quarters since Basel I with 0 corresponding to 1989Q1. 95% confidence intervals are shown and standard errors are two-way clustered at the bank and time levels.

B Data construction

B.1 Steps for variable selection

1. Consider only variables for total activity, i.e. remove variables pertaining to only domestic or foreign activity. Given that many banks only operate in the US, domestic activity equals total activity for most banks, and so including both domestic and total activity variables could generate collinearity.

2. Keep only those variables with values for at least 90% of observations in each quarter from 1984Q1 to 1999Q4.

3. Drop variables with a value of zero for 40% of observations in the sample.

4. Remove perfectly collinear variables and those manually identified to be highly collinear.

B.2 Variables used to estimate capital ratios

Table 3.3: Variables used to estimate capital ratios

Variable	**Notes**
Total assets	Sum of all asset items
Premises and fixed assets	Premises and fixed assets (including capitalised leases)
Cash and due	Total cash and balances due from depository institutions (including interest- & noninterest-bearing deposits)
Noninterest-bearing cash & due	Noninterest-bearing cash (currency and coin) and balances due from depository institutions
Cash items	Cash items in process of collection, unposted debits and currency and coin
Total deposits	Total deposits
Interest-bearing deposits	Interest-bearing consolidated office deposits
Deposit interest expense	Interest expense on total deposits (domestic and foreign)
Total interest expense	Total interest expense

All other non-interest expense	Includes all operating expenses of the entity not required to be reported elsewhere in regulatory returns
Premises and fixed assets expense	Expenses of premises and fixed assets (net of rental income and excluding salaries and employee benefits and mortgage interest)
Total bank equity capital	Total bank equity capital
Common stock	Includes the aggregate par or stated value of common stock issued
Surplus	Surplus (excludes all surplus related to preferred stock)
Salaries and employee benefits	Salaries and employee benefits
Income (loss) before income taxes and discontinued operations	Income (loss) before income taxes and discontinued operations
Loan income	Total interest and fees on loans
Interest income	Total interest income
Other noninterest income	All other noninterest income
IRAs and Keogh plans deposits	Individual retirement accounts (IRAs) and Keogh plan accounts held in domestic offices
Security income	Total interest and dividend income on: US Treasury securities, US government agency and corporation obligations, securities issued by states and political subdivision in the US, other domestic debt securities, foreign debt securities, and equity securities
Allowances for loan and transfer risk	Allowance for loan and lease financing receivable losses and allocated transfer risk
CI loans	Commercial and industrial loans
Consumer loans	Loans to individuals for household, family and other personal expenditures
Other consumer and related plans	Other revolving credit plans and other consumer loans
Loans and leases	Loans and lease financing receivables including unearned income
Allowances for loan and leases	Allowance for loan and lease financing receivable losses
Real estate loans	Loans secured by real estate
Non-farm non-residential real estate loans	Non-farm non-residential properties secured by real estate held in domestic offices

Residential loans	Loans secured by 1-4 family residential properties held in domestic offices
Total N/C loans and leases	Total loans and lease financing receivables 90 days or more past due and nonaccrual
Net income	Net income attributable to the bank
Net noninterest expense	Net noninterest expense (noninterest expense minus noninterest income)
Noninterest income	Total noninterest income
Noninterest expense	Total noninterest expense
Time deposits over $100,000	Time deposits of $100,000 or more held in domestic offices
Other assets	Other assets
Income earned not collected	Income earned, not collected
Other liabilities	Other liabilities
Expenses accrued and unpaid	Total expenses accrued and unpaid (includes accrued interest payable and income taxes payable)
All other liabilities	Includes amounts that cannot be reported properly against other liability items
Other real estate owned	The book value (not to exceed fair value), less any accumulated depreciation of all real estate other than bank premises actually owned by the bank and its consolidated subsidiaries
Securities	Total securities: the sum of held-to-maturity securities at amortised cost, available-for-sale securities at fair value and equity securities with readily determinable fair values not held for trading on a consolidated basis
US Agency	Total US government agency and corporation obligations
Allowance for credit losses	Allowance for credit losses

Note: this table provides a list and description of all variables considered in the non-linear least squares estimation of capital ratios (Equation 3.1). Data used is raw data directly from the Statistics of Depository Institutions (SDI) compiled by the Federal Deposit Insurance Corporation (FDIC).

TABLE 3.4: Dependent variables used in difference-in-differences estimation

Variable	**Notes**
Capital ratio	Obtained using Equation 3.1 and winsorised at the 1^{st} and 99^{th} percentiles of each quarter
Total assets	Sum of all asset items
Total loans	Total loans and leases, net of unearned income
Consumer loans	Loans to individuals for household, family and other personal expenditures
CI loans	Commercial and industrial loans
Residential loans	Loans secured by 1-4 family residential properties held in domestic offices
Real estate loans	Loans secured by real estate other than residential loans

Note: this table provides descriptions of dependent variables used in the difference-in-differences estimation (Equation 3.2). Data used is raw data directly from the Statistics of Depository Institutions (SDI) compiled by the Federal Deposit Insurance Corporation (FDIC).

Chapter 4

Credit, Capital and Crises: a GDP-at-Risk approach

1 Introduction

What is the relationship between vulnerabilities associated with elevated debt and asset prices and downside risks to economic growth? Recent research has established a strong relationship between indicators of financial conditions derived from asset prices and downside risks to growth in the *near term* up to one year ahead (Adrian et al., 2019). In this paper, we augment this programme of research by considering a wider set of macroprudential indicators, including measures of credit, house prices, external imbalance, and banking system resilience – information routinely monitored by central banks. We find that these indicators have forecasting power over downside risks to economic growth over the medium-term, specifically 3 to 5 years ahead.

We first construct a novel cross-country panel dataset covering 16 advanced economies over the period 1980:Q4-2017:Q4. For each country, we collect information on credit-to-GDP ratios, house price growth, current account imbalances, and a fast-moving measure of financial conditions. We also construct a measure of banking sector leverage computed as tangible common equity ratios, which we obtain by aggregating individual bank balance sheet information in each country. This permits us to assess the impact of the substantial increase in capital requirements, and hence banks' capital, following the Global Financial Crisis on downside risks. We apply quantile regressions (Koenker and Bassett, 1978) to estimate the relationship between these indicators and the shape of the GDP growth distribution across our panel. Using local projections (Jordà, 2005), we explore how this relationship varies up to 20 quarters ahead, focusing on the 12-quarter horizon as a benchmark. Given implementation and transmission lags, this arguably is the relevant policy horizon for implementing macroprudential policy responses to address the impact of building vulnerabilities.[1]

We find significant relationships between each of the vulnerability metrics and the 5$^{\text{th}}$ percentile of the future GDP growth distribution (which we refer to as "GDP-at-Risk").[2] Moreover, these relationships are both economically intuitive and meaningful in magnitude. Forecasting 12 quarters ahead, we find that GDP-at-Risk cumulatively deteriorates by 0.9, 0.75 and 1.5 percentage points following one-standard-deviation increases in the 3-year change in the credit-to-GDP ratio, 3-year real house price growth and the current account deficit (as a proportion of GDP) respectively. These results are consistent with findings from the early-warning literature that analyses the precursors of banking and currency crises (e.g., Reinhart and Kaminsky, 1999; Schularick and Taylor, 2012).

In a novel result, we find that higher bank capital mitigates these increases in risk: a one-standard-deviation increase in bank capitalisation, as measured by tangible common equity ratios, leads to a cumulative 0.9 percentage point improvement in GDP-at-Risk over three years. By contrast, the median projection does not significantly change in response to higher bank capital. This finding is consistent with theories that emphasise the role of bank capital as a buffer to absorb losses in a stress. Franta and Gambacorta (2020) provide collaborating evidence on the positive and significant role of macroprudential policies, in the context of loan-to-value ratio and loan provisions, in mitigating the risks to output growth. Similarly, Galán (2020) shows the benefits of macroprudential policies on the left-hand tail of GDP growth distribution. Boyarchenko et al. (2020) confirms the benefits of the increases in capital ratio in reducing downside risks to growth in the US. The positive impact of tighter macroprudential policy in mitigating the downside risks to growth originating from loose financial conditions is explored by Brandao-Marques et al. (2020).

In contrast to Adrian et al. (2018), we find no impact on 3-year-ahead GDP-at-Risk from movements in financial conditions or asset price volatility. The impact of these indicators is apparent only in the near term (i.e. at horizons of up to one year), over which time a tightening in financial conditions depresses GDP-at-Risk. This finding is in

[1] For instance, unless in exceptional circumstances, the countercyclical capital buffer has an implementation lag of one year. Moreover, macroprudential authorities may prefer to vary their countercyclical tools in a gradual manner (see, for example, Bank of England, 2016).

[2] See Cecchetti (2006) and De Nicolò and Lucchetta (2012) for early expositions of this approach, and Adrian et al. (2018) and Adrian et al. (2019) for more recent contributions to this literature.

line with evidence from Plagborg-Møller et al. (2020) that financial variables have limited forecasting power. Our findings are robust to alternative specifications of our regression equation such as the inclusion of the Miranda-Agrippino and Rey (2015) measure of the global financial cycle and single variable quantile regression setups.

Using our estimates, we illustrate the significant time variation in medium-term tail risks in advanced economies over the past four decades, decomposing the contributions of each of our vulnerability indicators. In the United States, our estimates point to a sharp deterioration in the 3-year-ahead forecast of GDP-at-Risk prior to both the early 1990s recession and the Global Financial Crisis driven by rapid growth in credit and house prices and, on the latter occasion, a widening current account deficit.

While this retrospective analysis is encouraging, we find that including the crisis episode and its aftermath is key to uncovering the impact of bank leverage on tail risk in our sample. When calculated over subsamples, we find an unstable relationship between these variables prior to 2007. This finding is perhaps unsurprising given that the Global Financial Crisis was the first simultaneous full-blown banking crisis hitting advanced economies since the Great Depression. More promisingly, the relationships between other vulnerability metrics and GDP tail risk are robust across subsamples. In particular, estimates of the impact of house prices, current account deficits and financial conditions remain stable. While there is some instability in the estimated impact of credit growth in our full baseline model, we find the impact of this indicator to be stable in single variable regressions.

These findings may be of interest to policymakers in central banks and other policy institutions charged with monitoring systemic risks in the financial system. Since the crisis, a plethora of such macroprudential frameworks and associated policy committees have been set up for this purpose. Edge and Liang (2019) document that such committees now exist in 47 countries around the world. A key challenge in operationalising these frameworks is improving our understanding of the impact of indicators of underlying vulnerabilities observable today on the potential for destabilising financial instability in future. Our findings contribute to our collective understanding of these relationships, and hence can inform the inferences policymakers draw from developments in different macroprudential indicators. They suggest the potential for conditioning the stance of macroprudential policy on such vulnerability indicators. These findings will also be of interest to researchers working to develop macroeconomic models that can generate crisis dynamics (Adrian and Boyarchenko, 2012; Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2014). Our results can inform the development and calibration of these models by providing some basic empirical facts about the precursors of tail risk events.

Our paper relates to three main strands of the literature: first, and most directly, we build on a strand of studies that use quantile regressions to estimate the distribution of GDP growth conditional on financial and economic conditions (Adrian et al., 2018; Aikman et al., 2018; Adrian et al., 2019; Franta and Gambacorta, 2020; Galán, 2020).[3] We contribute to this body of work by exploring how downside risk changes with respect to multiple indicators, including the effect of measures of banking system resilience.

Second, our work relates to the large literature on early warning indicators of financial crises, which seeks to find empirical regularities in the run-up to financial crises. Perhaps the most important result in this literature is the importance of credit-based variables as leading indicators of both the likelihood and severity of crises (see e.g., Schularick and Taylor, 2012; Jordà et al., 2013).[4] In this paper, we provide new evidence on the relationship between banking system capital ratios and macroeconomic tail risk. The closest empirical work to ours is Jordà et al. (2017), who examine the relationship between bank capital ratios and the probability and severity of crises using a large cross-country data set. While they find no relationship between measures of bank capital and the probability of crises, they show that conditional on being in a crisis, countries with better capitalised banking systems experience faster recoveries. While our procedure does not condition on crisis states, our results are qualitatively consistent with theirs in that we find that higher capital ratios improve tail growth outcomes over the medium term. Our finding is also consistent with microeconometric evidence that banks that entered the financial crisis with higher capital ratios contracted their lending by less (Carlson et al., 2013) and with work documenting the transmission of bank distress to real economic activity (see, for example, Chodorow-Reich (2014), who shows that bank distress led to an economically significant reduction in employment at small and medium-sized US firms reliant on bank credit). Third, our work relates to the growing literature on the real effects of macroprudential policy actions (e.g., International Monetary Fund, 2011; Kuttner and Shim, 2016; Bruno et al., 2017; Akinci and Olmstead-Rumsey, 2018; Richter et al., 2018).

The rest of the paper is organised as follows: Section 2 introduces our data and Section 3 describes our quantile regression methodology. Section 4 presents our results, while

[3] Previously, the impact of housing and equity price booms on tail risks were explored by Cecchetti (2006) and Cecchetti and Li (2008). Similarly, Giglio et al. (2016) employs quantile regressions to assess the predictive power of various systemic risk indicators.

[4] For research on the relationship between credit growth and financial crisis risk, see Gavin and Hausmann, 1996; McKinnon and Pill, 1996; Eichengreen and Arteta, 2000; Honohan, 2000; Bordo et al., 2001; Borio and Lowe (2002a,b, 2004); Borio and Drehmann, 2009; Drehmann et al., 2011; Mendoza and Terrones, 2014; Baron and Xiong, 2017; Bridges et al., 2017.

Section 5 concludes. Additional analysis and details of the dataset are left to Appendices A and B, respectively.

2 Data

Our analysis is based on a cross-country panel dataset using time series from 16 advanced economies over the period 1980:Q4-2017:Q4. These countries are: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom, and the United States.[5]

For each country, we collect time series for five vulnerability measures: *i)* the 3-year percentage point change in the private non-financial sector credit-to-GDP ratio; *ii)* 3-year real house price growth; *iii)* the current account deficit as a percentage of GDP; *iv)* realised volatility over one quarter in equity prices (we also report results replacing this with a financial conditions index); *v)* banking system tangible common equity (TCE) to total asset ratios as a measure of the resilience of the financial system. The TCE ratio is a widely-used measure of banks' resilience (see Basel Committee on Banking Supervision, 2010; Demirgüç-Kunt et al., 2013).[6] The measurement of indicators *i)* - *iii)* is relatively standard, but *iv)* and *v)* warrant some further discussion.

Bank capital

To construct a cross-country dataset for the TCE ratio, we first collect individual bank balance sheet data on group-level TCE (defined as common equity minus preference shares and intangible assets) and total tangible assets for banks in each of the aforementioned countries.[7] This information is obtained from Thomson Reuters Worldscope.[8] The TCE ratio for a bank is the ratio of its tangible common equity to tangible assets. To aggregate these data into a single country-level TCE ratio that is comparable over time, we use a chain-weighted approach, which allows us to take into account the entry and exit of banks each period. Details of this approach are provided in Appendix B with summary statistics

[5]We experimented with including Japan in this sample, but found that its inclusion generated implausibly large moves in some of the estimated coefficients. We re-ran our estimation removing each country individually, and the results did not change significantly when any other country was removed.

[6]The TCE measure we use is strongly correlated with other measures of banking system leverage. For instance, it has a correlation of 0.75 with the Bank of England's leverage indicator for the United Kingdom.

[7]Total assets here covers total cash and due from banks, investments, net loans, customer liability on acceptances, investment in unconsolidated subsidiaries, real estate assets, net property, plant and equipment and other assets.

[8]In general, Worldscope targets publicly quoted companies, and its coverage depends on certain criteria being met, such as a market capitalisation of over $100m or belonging to one of the major stock indices.

on the banks in our sample provided in Table 4.4. Data are available at annual frequency – our measure for year t is taken at the end of year t, and is linearly interpolated to create a quarterly series. As we discuss later, our results do not change significantly if we use the annual series.

Financial conditions

To estimate the impact of country-specific financial conditions, we explore two alternative variables. We first use equity price volatility as a proxy for financial conditions in our baseline specification to make use of its longer data availability. This series can be extended back to 1980 with the other variables in our specification. The volatility series is measured as the monthly standard deviation of daily returns in each country's equity price index. For robustness, we also show results using a financial conditions index (FCI) with a sample beginning in 1991 as in Eguren-Martin and Sokol (2020). This FCI is a modified version of that constructed by International Monetary Fund (2017), which follows the methodology of Koop and Korobilis (2014). The headline FCIs comprise of term spreads, interbank spreads, corporate spreads, sovereign spreads, long-term interest rates, policy rates, equity returns, and equity volatility. House price and credit growth variables are removed as they are introduced to the specification separately to isolate their impact.[9] The FCI and equity volatility series are strongly correlated; for the US, the correlation is 0.92, while for the UK it is 0.72.

It is a difficult task to disentangle the effect of each variable has on the complex financial system individually. The variables we use in our reduced-form analysis allow us to explore the correlation, rather than the causal relationship, of how downside risks to output growth change when a variety of indicators is considered. These variables exhibit correlations with each other with some being significant. However, these correlations are small, with the largest being 0.34, as Figure 4.11 demonstrates. This gives us confidence to consider multiple indicators in the same system.

In addition to the variables we discussed, we also use the central bank's policy rate and inflation rate alongside lagged quarterly GDP growth for each country in the empirical analysis as macroeconomic control variables. All variables are standardised by their country-level means and standard deviations. We provide details of the data sources and descriptive statistics in Appendix B.

[9]We would like to thank Fernando Eguren-Martin for providing these data. Eguren-Martin and Sokol (2020) discuss the properties of a related FCI measure and its global component.

3 Quantile regression methodology

In this section, we turn to quantile regressions to explore how the full distribution of real GDP growth varies with the vulnerability metrics described in the preceding section. Quantile regression is a widely-used technique that allows researchers to analyse how changes in a set of conditioning variables influence the shape of the distribution of the variable of interest (Koenker and Bassett (1978)). In our application, we estimate quantile regressions for a panel of advanced economy countries, requiring the treatment of country-specific fixed effects to avoid estimation bias. We follow Canay (2011) and assume that country fixed effects are locational shifts for the entire distribution (i.e. country fixed effects are the same across different quantiles). Under this assumption, we are able to employ a two-step procedure to eliminate country fixed effects and estimate our coefficients of interest.[10]

The first stage involves using a standard within estimator to estimate the fixed effects. We estimate the following linear pooled panel model by OLS:

$$y_{i,t+h} = \alpha_i^h + \gamma^h X_{i,t} + \epsilon_{i,t}, \tag{4.1}$$

The left-hand-side of Equation 4.1 is the average annualised growth rate of real GDP over h quarters, $y_{i,t+h}$, where $y_{i,t+h} = \frac{(Y_{i,t+h} - Y_{i,t})}{h/4}$ and $Y_{i,t+h}$ denotes the *log* level of real GDP of country i at time $t+h$ for horizons $h = 1, 2, \ldots, 20$ quarters. Our coefficient units are thus comparable across horizons. Fixed effects are denoted by α_i^h and X_{it} contains our vulnerability metrics and control variables described in Section 2 for country i measured at time t.[11]

Canay (2011) shows that the fixed effects can be estimated as:

$$\hat{\alpha}_i^h = \frac{1}{N} \sum_{i,t} (y_{i,t+h} - \hat{\gamma}^h X_{i,t})$$

In the second stage, we define the dependent variable as $y^*_{i,t+h} = y_{i,t+h} - \hat{\alpha}_i^h$. We then

[10] There are other ways of treating fixed effects in quantile regression setting, e.g., Galvao (2011). However, these methods rely on larger panel datasets to estimate fixed effects accurately at each quantile.

[11] In our baseline model, the y variable is not standardised which means that coefficients can be interpreted as percentage point changes in real GDP growth. The results do not change significantly if we standardise GDP growth as well as the explanatory variables.

proceed with quantile regressions as follows to estimate β_τ^h,

$$\hat{\beta}_\tau^h = \underset{\beta^h}{\operatorname{argmin}} \sum_{i,t} \rho_\tau(y_{i,t+h}^* - X_{i,t}\beta_\tau^h)$$

where τ denotes the quantile under consideration and ρ_τ is the standard asymmetric absolute loss function: $\rho_\tau(u) = u \times (\tau - \mathbb{1}\{u < 0\})$. The model is estimated from 1 to 20 quarters ahead using local projections (Jordà, 2005) to understand how the left tail of GDP growth develops over the forecast horizon. For inference, we follow the block bootstrapping method of Kapetanios (2008).[12] This method resamples the data over blocks of different time series dimensions to generate the standard errors of the estimated coefficients for respective quantiles. In our application, we resample the time series observations with replacement using 8 blocks (corresponds to 2 years), although changing the block size to 4 or 12 blocks does not alter our results.

4 Results

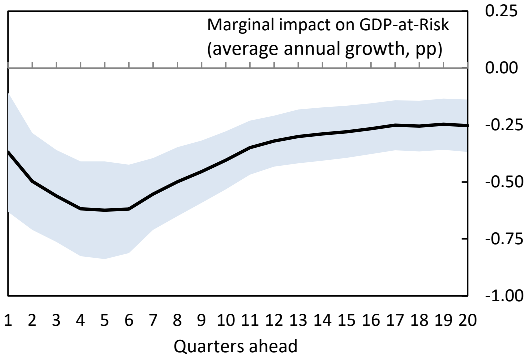
We first focus on the relationship between our vulnerability indicators and the projected 5$^{\text{th}}$ percentile of GDP growth (henceforth referred to as "GDP-at-Risk"). Figure 4.1 plots local projections of the estimated change in GDP-at-Risk at various horizons. The results are reported for common annualised GDP growth units. Note that we invert the sign of the current account balance and equity volatility following our priors that an increase in the current account deficit and periods of low volatility may bring about a deterioration in GDP-at-Risk over the medium term. Before turning to the results, it is important to emphasise that we have also performed the quantile regressions by considering one variable at a time, rather than using all of them in the same equation, and the local projection results remain broadly the same. This allows us to pursue the decomposition exercise in the next subsection.

Overall, the coefficients for credit and the current account are always negative. Therefore, stronger increases in credit-to-GDP ratios or a wider current account deficit has a detrimental effect on tail risk across our entire forecast horizon. Stronger house price growth appears to have a beneficial effect in the short term, but in the medium term this effect is more than offset and the coefficient is negative after around two years. The fast moving volatility measure is only significant in the short term, indicating that a sharp spike in this indicator increases tail risk immediately, but has little impact in the medium term.

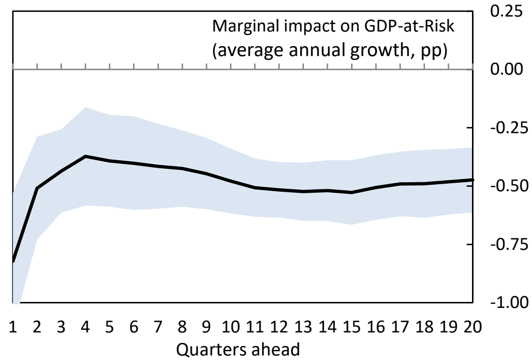
[12] See also Lahiri (2003).

Figure 4.1: Baseline results - impact on 5^{th} percentile of GDP growth at different horizons

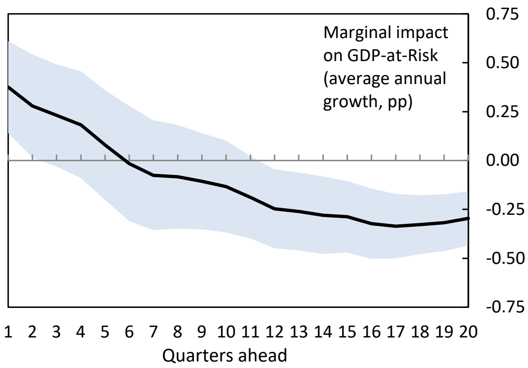
(a) Credit-to-GDP (3 year pp change)



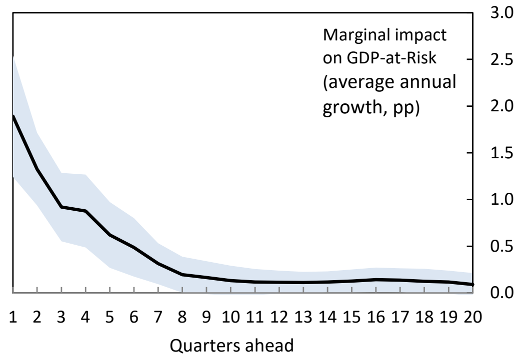
(b) Current account deficit



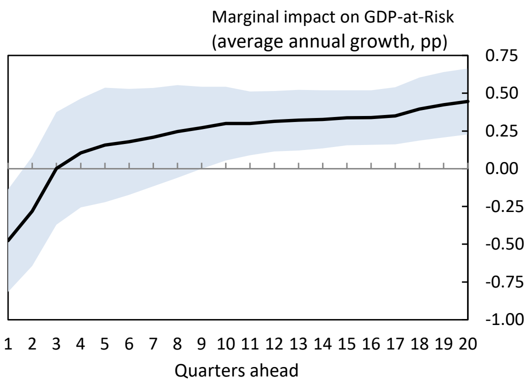
(c) Real house price growth (3 year)



(d) Volatility



(e) Bank capital (TCE) ratio



Note: These charts show the impact of a one-standard-deviation increase in a given indicator at time t on the 5^{th} percentile of real GDP growth at each horizon on the x-axis. GDP growth is measured as the average annual growth rate at each horizon. Confidence intervals represent ± 1 standard deviation. Standard errors are generated using block bootstrapping following Kapetanios (2008).

Finally, an increase in the capital ratio has a beneficial effect for GDP-at-Risk in the medium term. Our baseline specification also includes an intercept and controls, results for which are reported in Figure 4.7.

We proceed by discussing these results in two stages: first, we focus on the impact of innovations in vulnerabilities on GDP-at-Risk over the medium term, which we take as a three-year horizon. Given that the local projections presented in Figure 4.1 are relatively flat between quarters 12 and 20, our focus on the 12th quarter is representative of a broader medium-term (3-to-5 year) horizon.[13] Second, we discuss our results across the GDP growth distribution, expanding our attention beyond the 5th percentile GDP-at-Risk measure.

4.1 Downside risks to growth over the medium term

In Figure 4.2, we summarise the impact of each of our vulnerability indicators and macroeconomic controls on GDP-at-Risk at the three-year horizon. We discuss each indicator in turn.

Credit, house prices, and current account deficits

We find that medium-term tail risks to growth are aggravated by periods of rapid credit growth, house price growth, and large current account deficits. This chimes with insights from the voluminous literature on early warning indicators of financial crises, a typical finding of which is that credit booms accompanied by rapid house price inflation tend to increase the probability and severity of crises (see, for example, Kaminsky and Reinhart, 1999; Schularick and Taylor, 2012; Jordà et al., 2013; Aikman et al., 2018).

The estimated impacts of each of these three vulnerabilities on GDP-at-Risk are both statistically and economically significant. For example, a one-standard-deviation increase in the 3-year change of the credit-to-GDP ratio is associated with a 0.3 percentage point weaker GDP-at-Risk per annum over the next 3 years, thus cumulating to 0.9 percentage points over this period.[14] To give a sense of scale, between 2004 and 2007, the UK's credit-to-GDP ratio rose by 23 percentage points, 1.3 standard deviations above the mean change over the sample. Our credit result thus suggests that this was associated with a cumulative 1.2 percentage point deterioration in 3-year-ahead GDP-at-Risk over this

[13] Note that the local projections in Figure 4.1 give the average annual growth impact at each horizon. A flat, non-zero projection therefore implies a building cumulative level effect over time. For example, a coefficient of 0.25pp at the 4-year (16-quarter) horizon implies a total level effect of 1pp on GDP-at-risk. At the 5-year horizon it would imply a 1.25pp cumulative effect. If, instead, the level effect were permanent at 1pp, we would expect to see the projection gradually decay at longer horizons (to 0.2 in year 5, 0.17 in year 6, 0.14 in year 7, and so on).

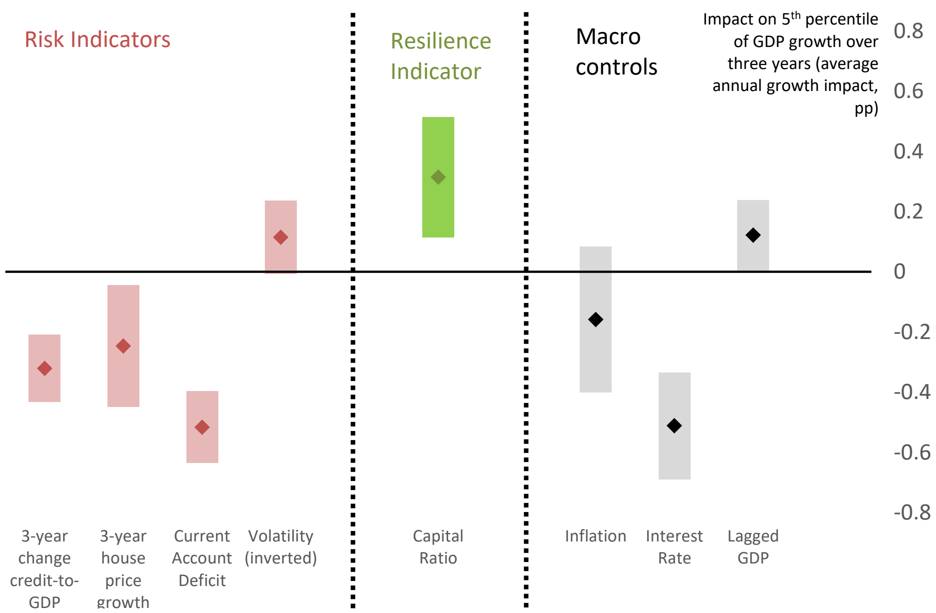
[14] As a robustness check, Figure 4.8 reports results of our baseline specification with credit split into its contributions from household and corporate borrowers. We find that after 20 quarters the effect of the changes in household credit is twice as severe as that of corporate credit.

period. The results remain the same if we use credit gap instead of credit-to-GDP ratio in the quantile regressions.

The estimated coefficient on real house price growth is similar in magnitude (-0.75 percentage points cumulatively), but somewhat less precisely estimated. The estimated impact of current account deficits on tail risk is twice as large, with a one-standard-deviation increase in the deficit increasing the severity of GDP-at-Risk in the medium term by 1.5 percentage points cumulatively. This is qualitatively consistent with potential amplification mechanisms associated with a heavy reliance on foreign funding. For example, to the extent that foreign flows prove relatively flighty, a large deficit may be associated with greater amplification of asset price and funding cost adjustments in the event of an adverse shock.

As a cross-check on these results, Appendix A reports results from an alternative specification of quantile regressions where the impact of each vulnerability indicator is estimated individually (see Figure 4.9).[15] We obtain broadly similar results in this exercise. The medium-term coefficients for real house price growth and the current account change very little, but the magnitude of the coefficient for credit increases by two-thirds.

FIGURE 4.2: Impact of each variable on 5^{th} percentile of GDP growth at 3-year horizon



Note: This figure shows the impact of a one-standard-deviation increase in a given indicator at time t on the 5^{th} percentile of real GDP growth after 12 quarters. The impact on GDP growth is measured as the average annual growth rate over 3 years. Confidence intervals represent a ± 1 standard deviation. Standard errors are generated using block bootstrapping following Kapetanios (2008).

[15] These regressions with individual vulnerability indicators also include macroeconomic controls.

Volatility and financial conditions

We find that a reduction in volatility is associated with a small decrease in the severity of GDP-at-Risk three years ahead. However this relationship is not statistically significant. As a cross-check on this finding, Table 4.1 (column 2) reports results from a regression where we replace our volatility measure with an index of financial conditions from Eguren-Martin and Sokol (2020).[16] Due to the availability of the index, we start our sample in 1991. Reassuringly, our baseline results do not materially change in this variant, and we continue to find only a small relationship between financial conditions and medium-term GDP-at-Risk.[17]

Adrian et al. (2018) show that loose financial conditions create an intertemporal trade-off in that they reduce tail risks in the near term at the expense of a modest deterioration in GDP-at-Risk in the medium term. We observe very similar results when our regression specification is stripped down to include just the financial conditions index and lagged GDP growth. However, the medium-term impact on GDP-at-Risk cannot be distinguished from zero when we add our various vulnerability indicators and further macroeconomic controls, with the change in the policy rate having a noticeable impact.[18] This finding is in line with evidence from Plagborg-Møller et al. (2020) that financial variables might have limited power on forecasting downside risks.

To the extent that the transmission of loose financial conditions to larger macroeconomic tail risks operates via boosting property prices and fostering excessive credit growth, we capture these channels directly with the inclusion of these variables. Indeed, Adrian et al. (2018) find that the impact of loose financial conditions on GDP-at-Risk in the medium term is amplified in the event of credit boom, defined as a dummy variable when credit growth is in the top 30 percent of its distribution. For the purposes of informing the gradual application of countercyclical macroprudential policy, our preferred approach is to estimate a continuous mapping from building credit vulnerabilities to GDP-at-Risk directly rather than relying on a binary credit boom indicator.

Given that changes in downside risks may be driven by global developments, we consider how fluctuations in the global financial cycle influence GDP-at-Risk. Our hypothesis is

[16] Eguren-Martin and Sokol (2020) follow the same methodology that the IMF employ in constructing cross-country FCIs which they regularly publish in their Global Financial Stability Reports, see e.g., https://www.imf.org/en/Publications/GFSR/Issues/2017/03/30/global-financial-stability-report-april-2017.

[17] An exception is the coefficient on real house price growth, which loses significance in this shorter sample.

[18] Adrian et al. (2018) include credit growth and house price measures within their FCI measure. In contrast, we strip these out of our FCI measure to avoid overlap with our slow-moving credit and house price vulnerability measures.

that when risk appetite is heightened globally, downside risks to growth over the medium term are more severe than if this is only a domestic development.[19] We explore this in the third column of Table 4.1 by re-estimating our baseline model with the global factor of Miranda-Agrippino and Rey (2015) replacing domestic equity volatility.[20]

As reported in Table 4.1, this global factor is found to have a material impact on GDP-at-Risk at the 3-year horizon. An increase in global asset prices (i.e. a loosening in global financial conditions) is estimated to increase the severity of a downturn by about 2 percentage points cumulatively over this horizon. This is consistent with Eguren-Martin and Sokol (2020), who find an important role for the global factor in their FCI measure. The coefficients on the other variables in our regression are broadly unaffected by the inclusion of a global factor: the coefficients on credit and the current account are of a similar magnitude, and the coefficients on house prices and capital have the same sign, but a smaller size. Overall, this relative stability in our estimates indicates that the global factor provides additional information over our sample that is uncorrelated with our other regressors.

Bank capital

Turning to the impact of financial system resilience, we find that higher levels of banking system capital significantly improve GDP-at-Risk in the medium term. This is a novel finding, consistent with the notion that credit crunch amplification mechanisms are a key driver of severe macroeconomic tail events and that higher banking sector capitalisation can forestall these adverse dynamics. We find that a one-standard-deviation increase in the banking sector's TCE ratio improves GDP-at-Risk by 0.9 percentage points cumulatively over the following three years. As an illustration, the United Kingdom's TCE ratio averaged 4.1% over our full sample with a standard deviation of 0.9 percentage points. In 2007, this ratio had fallen to 1.9%, 2.5 standard deviations below its average level. We estimate that this diminution in resilience alone is sufficient to account for a 2.3 percentage point deterioration in GDP-at-Risk cumulatively from 2008 to 2010.

One potential concern is that our bank capital measure is based on annual bank reports and has been interpolated to a quarterly frequency in order to match the frequency of other series in our panel. When we repeat our analysis with annual data, we obtain a near-identical 0.3 percentage point coefficient on capital at the three-year horizon and the

[19] Alessi and Detken (2011) find measures of global liquidity to be amongst the best leading indicators of financial crises in OECD countries. Cesa-Bianchi et al. (2019) report a similar finding.

[20] The results are broadly unchanged in an alternative specification where the global factor is included in addition to domestic equity volatility.

coefficient remains statistically significant (see Table 4.1 column 4).[21]

4.1.1 Decomposing GDP-at-Risk

In Figure 4.3, we use our baseline regression results for the medium term (3 years ahead) as a lens through which to view the drivers of tail risks to growth in the United Kingdom and United States over our sample. The upper panel shows the time series of predicted UK GDP-at-Risk, while the lower panel shows the estimated series for the United States. The black solid line shows the level of tail risk 3 years after each point in time as predicted by our model. For example, the reading for 2005:Q1 is the 5$^{\text{th}}$ percentile of the distribution of average annual GDP growth over the period 2005:Q1-2008:Q1 as predicted in 2005:Q1. One important caveat to this exercise is that we do not identify orthogonal disturbances. Rather, the contributions in this case show the impact on the risk projection of "news" in the time series of each of the right-hand-side variables of our regression.

Our model suggests that medium-term tail risks to growth have fluctuated significantly in both countries over our sample period. In the United Kingdom, GDP-at-Risk reached highly elevated levels prior to the 1990-1991 recession, driven by rapid growth in credit and house prices, an expanding current account deficit and extremely tight monetary conditions following increases in Bank Rate from 7% in May 1988 to almost 15% in October 1989. Each of these factors went into reverse following the recession, ushering in a prolonged period where risks to growth were subdued.

This benign period continued up until the late 1990s/early 2000s, when rapid growth in credit and house prices resumed, this time accompanied by weaker bank capital adequacy. This created a large and persistent increase in growth tail risks by the mid-2000s. By 2006:Q2, over two years before the failure of Lehman Brothers heralded the worst of the Global Financial Crisis, our model predicts that GDP-at-Risk was -3.9% cumulatively over the subsequent 3 years. In the aftermath of the crisis, our model views risks to the economy as having declined significantly, driven by modest increases in credit and house prices and the strengthening in banking system capital. The increase in bank capital is estimated to have reduced tail risks to growth by nearly 4 percentage points cumulatively. Offsetting these positive developments to some extent, however, has been the increasing current account deficit.

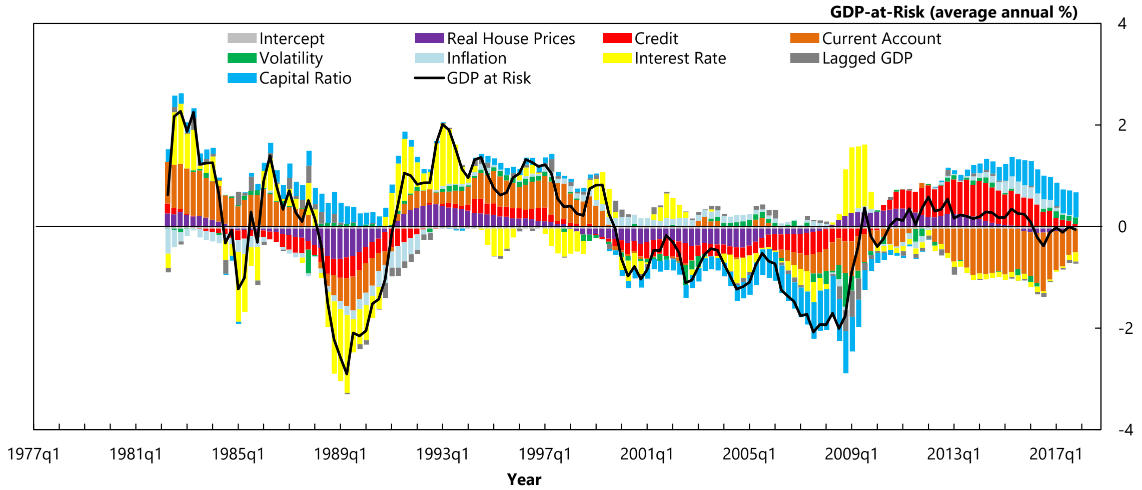
Our estimate of GDP-at-Risk for the United States shares a remarkably similar time

[21] We take end-year measures of our risk indicators and macroeconomic controls to match the frequency of the bank capital series.

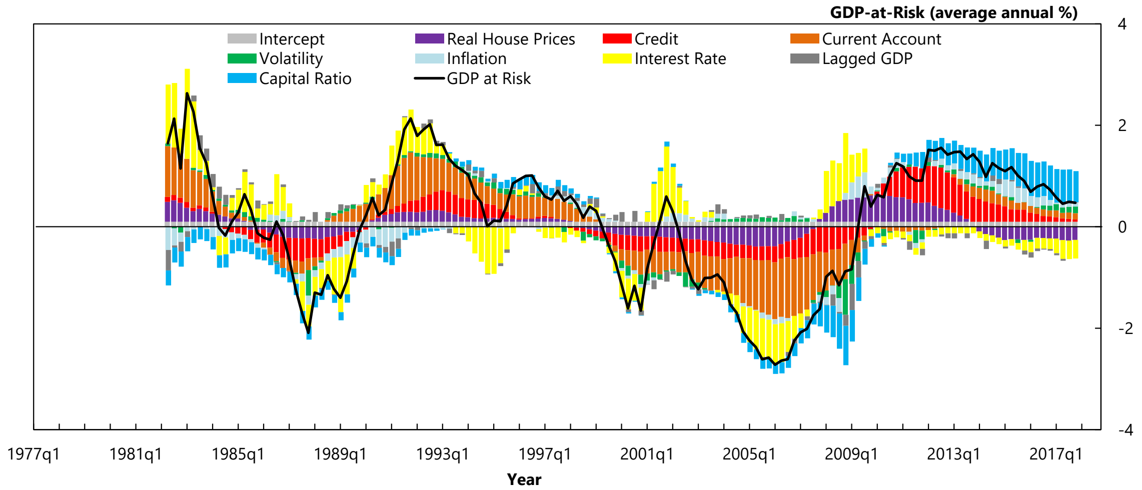
path. Risks to growth are estimated to have built significantly in the mid-to-late 1980s, driven by rapid growth in credit and house prices and against the backdrop of a weakly capitalised banking system. These risks were increased materially by the tightening in monetary policy in the late 1980s, culminating in the 1990-1991 recession. Just as for the United Kingdom, a benign period followed where tail risks to growth remained persistently subdued. Unsurprisingly, given the absence of equity valuations in our model, we miss the mild recession in 2001 that followed the collapse of the dot-com bubble.

FIGURE 4.3: Decomposition of GDP-at-Risk at the 3-year horizon

(A) UK – 3 years ahead



(B) USA – 3 years ahead



Note: The black solid line shows the average annual 5^{th} percentile of GDP growth 3 years after each point in time, as predicted by our model and using coefficients estimated from the full sample. The bars show the contribution of each indicator to that total. The cumulative impact at each point can be calculated by multiplying by 3.

We do, however, capture an unprecedented build-up in GDP-at-Risk from the mid-2000s onwards, driven by rapid growth in credit and house prices, and notably the widening in the current account deficit.[22] Many contemporaneous accounts emphasised risks associated

[22]In contrast to the United Kingdom, our measure of banking system capital does not contribute to

with the build-up in the US external deficit, which exceeded 6% of GDP in 2006. Our perspective, similar to Obstfeld and Rogoff (2009), is that the US current account deficit – and its counterpart, abundant inflows of capital to the US economy, intermediated by the financial system – was a strong signal of building internal imbalances over this period, which manifested themselves via an explosion in leverage in the shadow banking system and via a build-up in indebtedness in the household sector. By 2006:Q2, our model predicts that US GDP-at-Risk over the subsequent 3 years had reached -8% cumulatively. In the post-crisis period, we estimate that the severity of GDP-at-Risk has fallen substantially, driven to a large extent by the strengthening in banking system capitalisation, the slowing of credit growth and narrowing of the current account deficit.

4.1.2 Measuring GDP-at-Risk over subsamples

Figure 4.4 presents coefficients for GDP-at-Risk 3 years ahead, estimated using different sub-samples of our dataset. In particular, the far-left bar for each variable reports the 3-year-ahead coefficient estimate for the truncated sample period of right-hand-side variables observed from 1980:Q4 to 1992:Q1 (that is, including their impact on GDP realisations up to 1995:Q1); subsequent bars then expand the sample with an incremental 5 years of data. Figure 4.4a presents results using sub-samples of our full baseline model, while Figure 4.4b presents results using a simpler models only including each vulnerability indicator in turn (including controls).

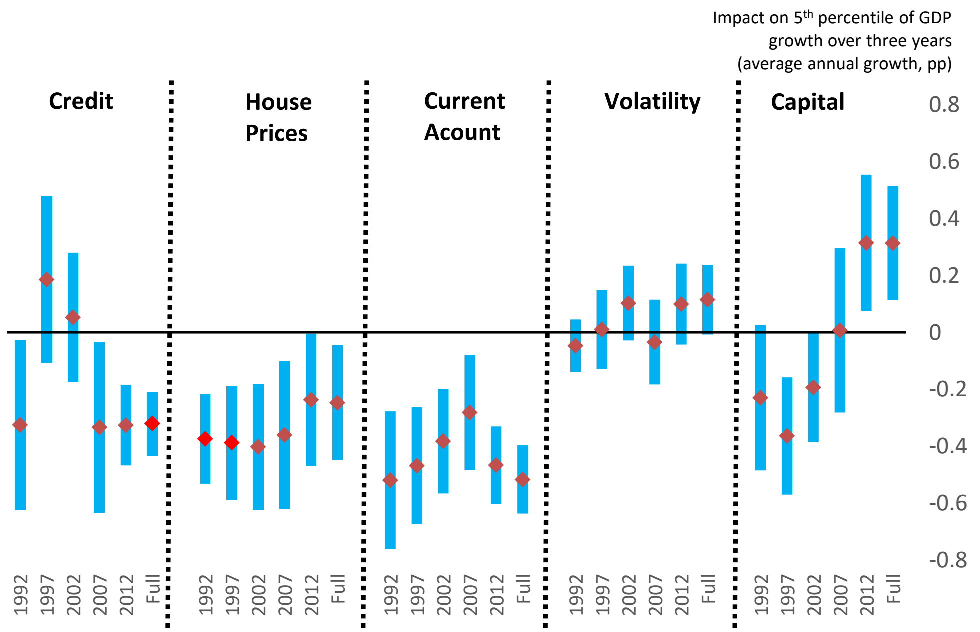
Overall, while the coefficient estimates for house prices, credit, current account deficits, and volatility are relatively stable over these sub-samples, the estimated impacts of bank capital can vary significantly, both in terms of magnitude and sign. In particular, a researcher estimating this regression in the early 2000s would have found a *negative* relationship between banking system capitalisation and GDP-at-Risk (i.e. more bank capital increases recession severity). This is perhaps unsurprising given that the Global Financial Crisis was the first simultaneous full-blown banking crisis hitting advanced economies since the Great Depression.

We offer two considerations for interpreting these results: first, the instability of our estimated capital coefficient emphasises the challenges involved in uncovering the impact of vulnerability metrics on extreme tails of the distribution of growth, using what remains

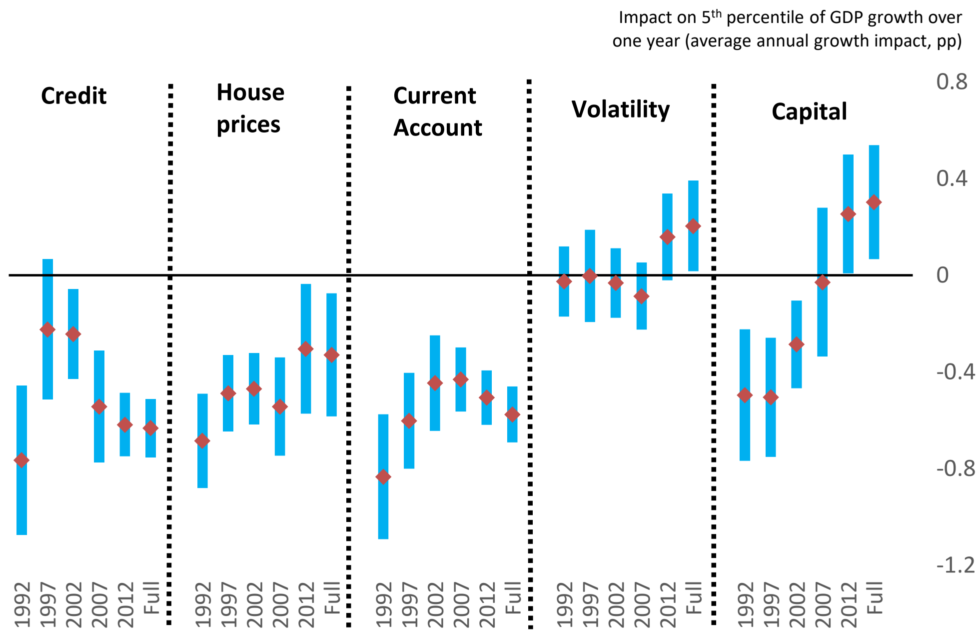
the deterioration in US GDP-at-Risk over this period. Commercial bank leverage, which our metric captures, was relatively stable over this period, with the increase in leverage concentrated in the large dealer institutions (Duffie, 2019).

Figure 4.4: Impact of each variable on 5th percentile of GDP growth at 3-year horizon over different sub-samples

(A) Full model



(B) Single variable model



Note: The figure shows how the 12 quarter coefficients in our baseline model (A) and a simpler model (B), which includes each variable individually (with macroeconomic controls), change if we restrict the vulnerabilities sample at each of the points on the x-axis.

a relatively small sample of data.[23] As such, caution is required when using results from such exercises to inform real-time risk assessment.[24] Second, it is plausible that having seen genuinely extreme observations in indicators and growth before and after the Global Financial Crisis, the 5th percentile coefficients in this regression will be less responsive to new data henceforth.

4.2 Characterising the full predicted GDP growth distribution

Our last set of results compares estimates of the tail of the predicted distribution of GDP growth with other parts of the distribution. We focus on comparisons with the 50th percentile (the median) and the 95th percentile. Figure 4.5 presents coefficient estimates for the 5th, 50th, and 95th percentiles, as well as OLS estimates, at the 3-year-ahead horizon. Our main finding here is that the impact of our vulnerability measures on growth is, by and large, estimated to have the same sign across all percentiles. It is notable that the current account loads more heavily on the left-hand tail in the medium term than on other parts of the distribution.

To illustrate the economic significance of these estimates, Figure 4.6a presents time series estimates of predicted percentiles of UK GDP growth 3 years ahead. The dotted lines shows the actual outturn of real GDP growth at each horizon. In order to aid comparison with actual outturns, we have shifted our GDP estimates forward relative to Figure 4.3. For example, the point labelled 2008 gives our forecast for 2008 GDP made three years ahead (in 2005).

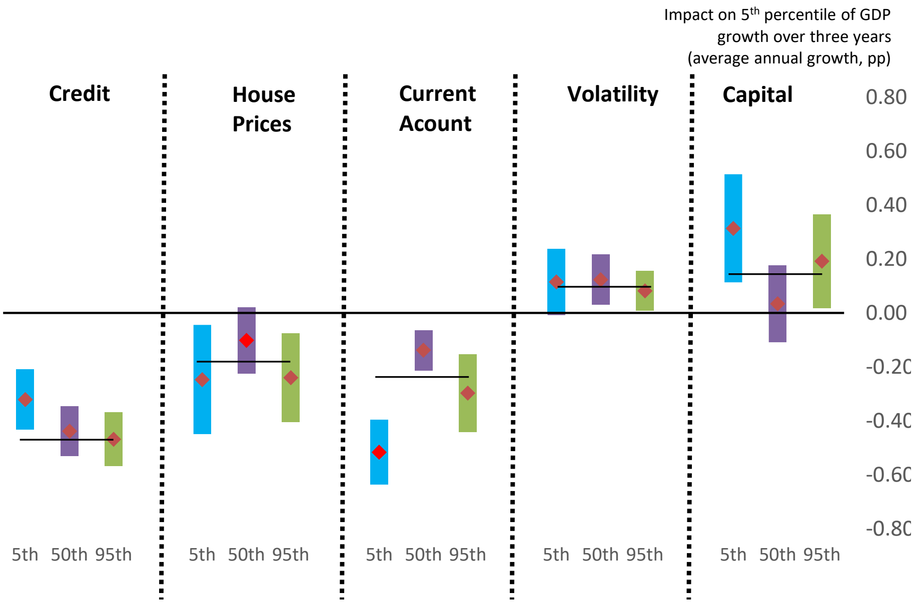
The outturns of GDP growth do not fall outside the lower 5% region of the predicted density. We find that innovations in vulnerability indicators act more like location shifters for the entire predicted density of GDP growth 3 years ahead, with both the 5th and 95th percentiles varying significantly (although the distance between these points of the distribution does increase in the run up to stress events).[25]

[23] This is reminiscent of the observation in Mendoza and Terrones (2014) in their analysis of credit booms, which updated an earlier analysis from 2008 with data from 2007-2010. The additional four years data had generated a *"a critical change from our previous findings because, lacking the substantial evidence from all the recent booms and crises, we had found only 9 percent frequency of banking crises after credit booms for emerging markets and zero for industrial countries."*.

[24] Challenges posed by real-time assessments of cyclical fluctuations are by no means unique to our approach or application. For example, real-time assessments of economic slack differ notably from such estimates made with the benefit of hindsight (e.g., Orphanides and van Norden, 2002; Edge and Rudd, 2016). This concern has also been emphasized in the literature on the credit-to-GDP gap (e.g., Edge and Meisenzahl, 2011).

[25] In Appendix A.3, we broaden the analysis in Figure 4.6a by calculating the 3-year-ahead forecast for GDP growth at *every* decile in the distribution, as well as at the 1st, 5th, 95th, and 99th percentiles. Figure 4.10 illustrates the proportion of actual GDP observations falling into each percentile bucket predicted by our baseline model, and shows that the fraction of observations falling into each part of the predicted

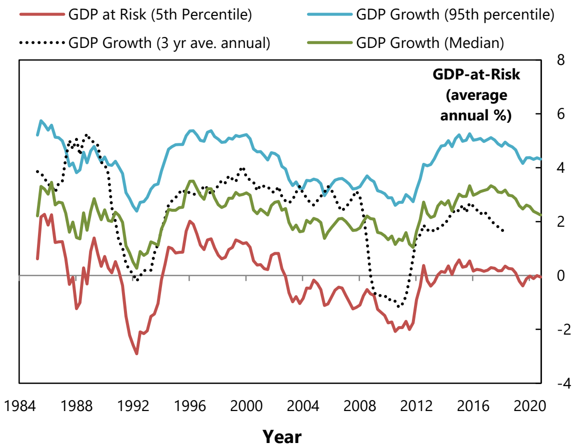
Figure 4.5: Impact of each variable on the 5th, 50th and 95th percentiles and conditional mean of GDP growth



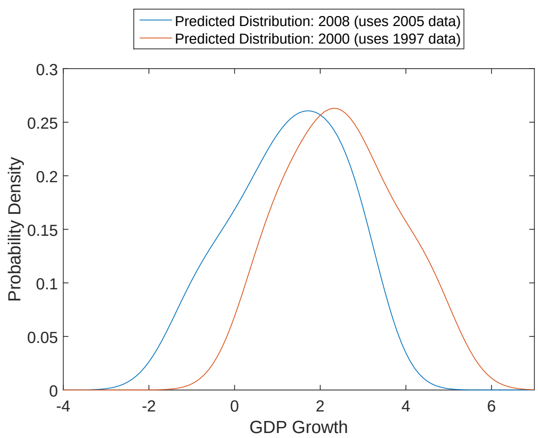
Note: This figure shows the impact of a one-standard-deviation increase in a given indicator at time t on a particular percentile of real GDP growth after 12 quarters. The OLS estimates are given by the horizontal black line for each indicator. Impact on GDP growth is measured as the average annual growth rate impact at the labelled percentile. Confidence intervals represent ± 1 standard deviation. Standard errors are generated using block bootstrapping following Kapetanios (2008).

Figure 4.6: Predicted GDP growth density

(a) Forecast from 3 years previously vs. actual outturn



(b) Predicted density (3 years ahead)



Note: The left panel shows the predicted 5th, 50th and 95th percentiles of GDP growth using data 12 quarters ahead of each point in time as well as the realised observations. The right panel shows the full predicted distribution of GDP growth in 2008 and 2000 using data from 2005 and 1997.

Finally, Figure 4.6b plot predicted densities of UK growth for 2008:Q3 as of 3 years

distribution are closely aligned with the expected proportions.

beforehand. These are obtained by applying a kernel density estimator to our full-sample quantile regression coefficients (estimated at the 5^{th} percentile, 95^{th} percentile, and every decile in between). Relative to a baseline predicted density for the year 2000 (shown for comparison), a researcher armed with this model in 2005:Q3 would have predicted a marked leftward shift in the entire distribution and a fattening in the left-hand tail, well in advance of the crisis that was to follow. These are retrospective estimates that rely on coefficients estimated using the full sample that would not have been obtainable at the time, and as such care should be taken in interpreting their utility for real-time risk assessment purposes.

5 Conclusion

The provision of sufficient early warning when downside risks to future growth increase is crucial for the successful operationalisation of the macroprudential frameworks that have been established worldwide as a legacy of the Global Financial Crisis. In this paper, we have developed a rich empirical framework, within which we trace the impact of a set of vulnerability measures on the real GDP growth distribution at various horizons. Our primary focus has been on the tail of the GDP distribution – GDP-at-Risk – and its determinants in the medium-term (at the 3-5 year horizon). Most importantly, we provide a framework within which a lack of financial system resilience is linked explicitly to downside risks to economic growth.

Drawing on our panel data across 16 advanced economies, we establish that familiar indicators of macrofinancial imbalance systematically increase GDP tail risks in the medium term. Credit booms, which have preceded around three-quarters of the worst GDP catastrophes in our sample, are found to materially worsen GDP-at-Risk in the medium term. We also find significant roles for rapid house price growth and a large current account deficit in affecting GDP tail risks three years out. We demonstrate that an increase in bank capital can improve GDP-at-Risk in the medium-term.

Our paper contributes to a programme of research that is required in order to deepen the evidence base underpinning macroprudential strategy. The framework we present could – and should – be extended in several dimensions: first, our set of vulnerability indicators is by no means exhaustive. Taking credit as an example, fruitful extensions include analysis of the relative roles of different types of credit (by sector or type of lender), the role of debt serviceability and the importance of the distribution of a given level of debt. The global nature of the financial cycle and the importance of international spillovers between our vulnerabilities should also be explored further. Moreover, our bank capital indicator is only one measure of financial system resilience and extensions to capture the role of liquidity both within the banking sector and in market-based finance are warranted.

A second dimension for future work is to establish structural counterparts to our empirical framework, which are able to generate the observed links between vulnerabilities and the GDP distribution. This would allow us to better understand the joint determination of our vulnerability indicators, thresholds above which they signal particular concern and to learn more about the underlying drivers of GDP-at-Risk.

Finally, we need to establish tools to better understand the transmission of macroprudential policy onto the GDP distribution. That transmission might operate directly – as in the link we have established from bank capital to GDP-at-Risk in this paper. Transmission may also operate indirectly, perhaps by leaning on the build-up of certain vulnerabilities or changing the extent to which a given aggregate imbalance transmits to risks at the borrower level. Assessing the transmission mechanism of different macroprudential tools through a common lens of their impact on the GDP distribution at different horizons would help to advance policy decisions on tool selection, the potential for tool interaction and the cost-benefit analysis critical for policy calibration.

References

Adrian, T., and N. Boyarchenko (2012): "Intermediary leverage cycles and financial stability," Staff Reports 567, Federal Reserve Bank of New York.

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable Growth," *American Economic Review*, 109(4), 1263–89.

Adrian, T., F. Grinberg, N. Liang, and S. Malik (2018): "The Term Structure of Growth-at-Risk," Working Paper 180, International Monetary Fund.

Aikman, D., J. Bridges, S. Burgess, R. Galletly, I. Levina, C. O'Neill, and A. Varadi (2018): "Measuring risks to UK financial stability," Bank of England working papers 738, Bank of England.

Akinci, O., and J. Olmstead-Rumsey (2018): "How effective are macroprudential policies? An empirical investigation," *Journal of Financial Intermediation*, 33, 33 – 57.

Alessi, L., and C. Detken (2011): "Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity," *European Journal of Political Economy*, 27(3), 520 – 533.

Bank of England (2016): "The Financial Policy Committee's approach to setting the countercyclical capital buffer," Policy statement, Bank of England.

Baron, M., and W. Xiong (2017): "Credit Expansion and Neglected Crash Risk," *The Quarterly Journal of Economics*, 132(2), 713–764.

Basel Committee on Banking Supervision (2010): "Calibrating regulatory minimum capital requirements and capital buffers: a top-down approach," Working paper, Bank of International Settlements.

Bordo, M., B. Eichengreen, D. Klingebiel, M. S. Martinez-Peria, and A. K. Rose (2001): "Is the Crisis Problem Growing More Severe?," *Economic Policy*, 16(32), 53–82.

Borio, C., and M. Drehmann (2009): "Assessing the risk of banking crises - revisited," BIS Quarterly Review March 2009, Bank of International Settlements.

Borio, C., and P. Lowe (2002a): "Assessing the risk of banking crises," BIS Quarterly Review December 2002, Bank of International Settlements.

——— (2002b): "Asset prices, financial and monetary stability: exploring the nexus," BIS Working Papers 114, Bank of International Settlements.

——— (2004): "Securing sustainable price stability: should credit come back from the wilderness?," BIS Working Papers 157, Bank of International Settlements.

Boyarchenko, N., D. Giannone, and A. Kovner (2020): "Bank Capital and Real GDP Growth," Staff Reports 950, Federal Reserve Bank of New York.

Brandao-Marques, L., R. G. Gelos, M. Narita, and E. Nier (2020): "Leaning Against the Wind: A Cost-Benefit Analysis for an Integrated Policy Framework," IMF Working Papers 2020/123, International Monetary Fund.

Bridges, J., C. Jackson, and D. McGregor (2017): "Down in the slumps: the role of credit in five decades of recessions," Bank of England working papers 659, Bank of England.

Brooke, M., O. Bush, R. Edwards, J. Ellis, B. Francis, R. Harimohan, K. Neiss, and C. Siegert (2015): "Measuring the macroeconomic costs and benefits of higher UK bank capital requirements," Financial Stability Paper 35, Bank of England.

Brunnermeier, M. K., and Y. Sannikov (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104(2), 379–421.

Bruno, V., I. Shim, and H. S. Shin (2017): "Comparative assessment of macroprudential policies," *Journal of Financial Stability*, 28, 183 – 202.

Canay, I. A. (2011): "A simple approach to quantile regression for panel data," *The Econometrics Journal*, 14(3), 368–386.

Carlson, M., H. Shan, and M. Warusawitharana (2013): "Capital ratios and bank lending: A matched bank approach," *Journal of Financial Intermediation*, 22(4), 663 – 687.

Cecchetti, S. G. (2006): "Measuring the Macroeconomic Risks Posed by Asset Price Booms," Working Paper 12542, National Bureau of Economic Research.

Cecchetti, S. G., and H. Li (2008): "Measuring the impact of asset price booms using quantile vector autoregressions," Mimeo, Brandeis University.

Cesa-Bianchi, A., F. E. Martin, and G. Thwaites (2019): "Foreign booms, domestic busts: The global dimension of banking crises," *Journal of Financial Intermediation*, 37, 58 – 74.

Chodorow-Reich, G. (2014): "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-09 Financial Crisis," *Quarterly Journal of Economics*, 129(1), 1–59.

De Nicolò, G., and M. Lucchetta (2012): "Systemic Real and Financial Risks : Measurement, Forecasting, and Stress Testing," *IMF Working Paper*, 58.

Demirgüç-Kunt, A., E. Detragiache, and O. Merrouche (2013): "Bank Capital: Lessons from the Financial Crisis," *Journal of Money, Credit and Banking*, 45(6), 1147–1164.

Drehmann, M., C. Borio, and K. Tsatsaronis (2011): "Anchoring countercyclical capital buffers: the role of credit aggregates," BIS Working Papers 355, Bank of International Settlements.

Duffie, D. (2019): "Prone to Fail: The Pre-crisis Financial System," *Journal of Economic Perspectives*, 33(1), 81–106.

Edge, R. M., and J. N. Liang (2019): "New Financial Stability Governance Structures and Central Banks," Finance and Economics Discussion Series 2019-019, Board of Governors of the Federal Reserve System (US).

Edge, R. M., and R. R. Meisenzahl (2011): "The Unreliability of Credit-to-GDP Ratio Gaps in Real Time: Implications for Countercyclical Capital Buffers," *International Journal of Central Banking*, 7(4), 261–298.

Edge, R. M., and J. B. Rudd (2016): "Real-Time Properties of the Federal Reserve's Output Gap," *The Review of Economics and Statistics*, 98(4), 785–791.

Eguren-Martin, F., and A. Sokol (2020): "Attention to the tail(s): Global financial conditions and exchange rate risks," ECB Working Paper 2387, Frankfurt a. M.

Eichengreen, B., and C. Arteta (2000): "Banking Crises in Emerging Markets: Presumptions and Evidence," Center for International and Development Economics Research (CIDER) Working Papers C00-115, UC Berkeley.

Franta, M., and L. Gambacorta (2020): "On the effects of macroprudential policies on Growth-at-Risk," *Economics Letters*, 196, 109501.

Galán, J. E. (2020): "The benefits are at the tail: Uncovering the impact of macroprudential policy on growth-at-risk," *Journal of Financial Stability*, p. 100831.

Galvao, A. F. (2011): "Quantile regression for dynamic panel data with fixed effects," *Journal of Econometrics*, 164(1), 142 – 157, Annals Issue on Forecasting.

Gavin, M., and R. Hausmann (1996): "The Roots of Banking Crises: The Macroeconomic Context," IDB Publications (Working Papers) 6067, Inter-American Development Bank.

Giglio, S., B. Kelly, and S. Pruitt (2016): "Systemic risk and the macroeconomy: An empirical evaluation," *Journal of Financial Economics*, 119(3), 457 – 471.

He, Z., and A. Krishnamurthy (2014): "A Macroeconomic Framework for Quantifying Systemic Risk," Working Paper 19885, National Bureau of Economic Research.

Honohan, P. (2000): "Banking System Failures in Developing and Transition Countries: Diagnosis and Prediction," *Economic Notes*, 29(1), 83–109.

International Monetary Fund (2011): "Macroprudential Policy; What Instruments and How to Use them? Lessons From Country Experiences," IMF Working Papers 11/238, International Monetary Fund.

——— (2017): "Global Financial Stability Report October 2017: Is Growth at Risk?," Report, International Monetary Fund.

Jordà, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95(1), 161–182.

Jordà, O., B. Richter, M. Schularick, and A. M. Taylor (2017): "Bank Capital Redux: Solvency, Liquidity, and Crisis," Working Paper 23287, National Bureau of Economic Research.

Jordà, O., M. Schularick, and A. M. Taylor (2013): "When Credit Bites Back," *Journal of Money, Credit and Banking*, 45(s2), 3–28.

Kaminsky, G. L., and C. M. Reinhart (1999): "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," *American Economic Review*, 89(3), 473–500.

Kapetanios, G. (2008): "A bootstrap procedure for panel data sets with many cross-sectional units," *The Econometrics Journal*, 11(2), 377–395.

Koenker, R., and G. Bassett (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

Koop, G., and D. Korobilis (2014): "A new index of financial conditions," *European Economic Review*, 71, 101 – 116.

Kuttner, K. N., and I. Shim (2016): "Can non-interest rate policies stabilize housing markets? Evidence from a panel of 57 economies," *Journal of Financial Stability*, 26, 31–44.

Lahiri, S. N. (2003): *Resampling Methods for Dependent Data*. Springer Series in Statistics.

McKinnon, R. I., and H. Pill (1996): "Credible liberalizations and international capital flows: the overborrowing syndrome," in *Financial Deregulation and Integration in East Asia*, ed. by T. Ito, and A. O. Krueger. Chicago University Press.

Mendoza, E. G., and M. E. Terrones (2014): "An Anatomy of Credit Booms and their Demise," in *Capital Mobility and Monetary Policy*, ed. by M. F. D., C. E. Raddatz, and C. M. Reinhart, vol. 18 of *Central Banking, Analysis, and Economic Policies Book Series*, chap. 6, pp. 165–204. Central Bank of Chile.

Miranda-Agrippino, S., and H. Rey (2015): "US Monetary Policy and the Global Financial Cycle," NBER Working Papers 21722, National Bureau of Economic Research, Inc.

Obstfeld, M., and K. Rogoff (2009): "Global Imbalances and the Financial Crisis: Products of Common Causes," CEPR Discussion Papers 7606, Centre for Economic Policy Research.

Orphanides, A., and S. van Norden (2002): "The Unreliability of Output-Gap Estimates in Real Time," 84(4), 569–583.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is growth at risk?," *Brookings Papers on Economic Activity,*, (4).

Reinhart, C. M., and G. L. Kaminsky (1999): "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," *American Economic Review*, 89(3), 473–500.

Richter, B., M. Schularick, and I. Shim (2018): "The macroeconomic effects of macroprudential policy," Working Paper 740, Bank of International Settlements.

Schularick, M., and A. M. Taylor (2012): "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008," *American Economic Review*, 102(2), 1029–61.

Appendix

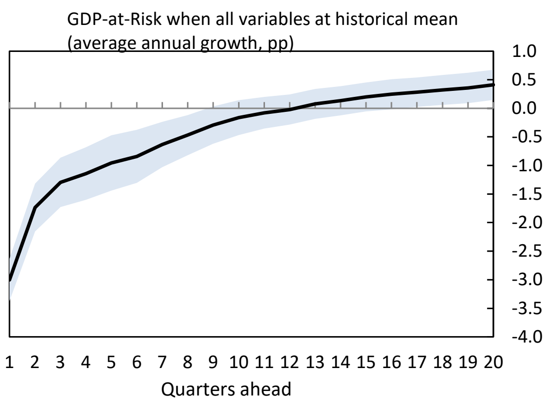
A Robustness checks and additional material

A.1 Results for the intercept and control variables in baseline model

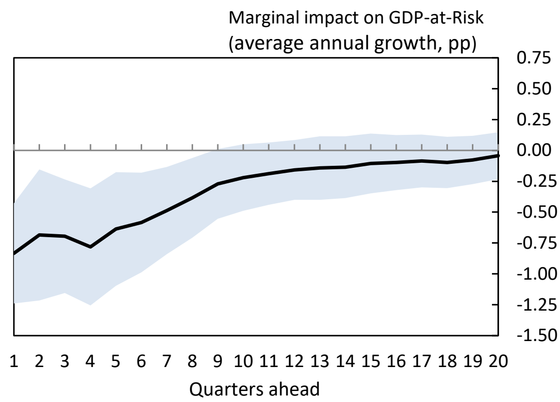
Figure 4.1 plots local projections of the estimated change in the GDP-at-Risk at various horizons, conditional on a one-standard-deviation change in each of the vulnerability indicators in our baseline model. In Figure 4.7, we report results for the intercept and control variables from the same specification.

FIGURE 4.7: Baseline results - 5$^{\text{th}}$ percentile: intercept and controls

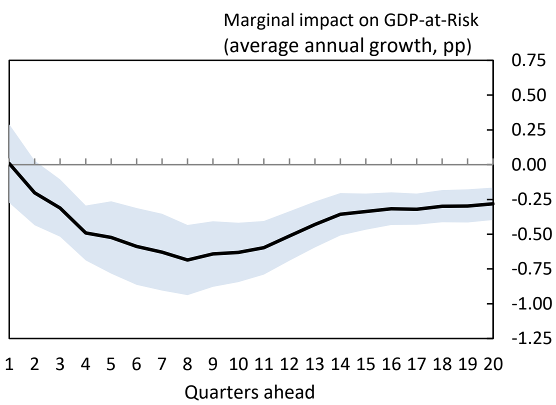
(A) Intercept



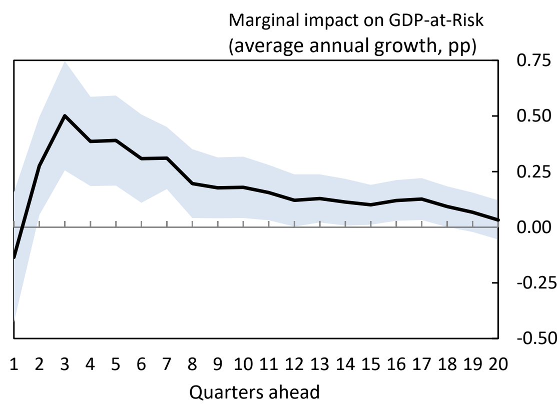
(B) Inflation



(C) Policy Rate



(D) Lagged GDP Growth



Note: Charts display coefficients for the intercept and control variables that were included in our baseline specification in Figure 4.1. Charts show the impact of a one-standard-deviation increase in a given indicator at time t on the 5$^{\text{th}}$ percentile of real GDP growth at each horizon on the x-axis. GDP growth is measured as the average annual growth rate at each horizon. Confidence intervals represent ± 1 standard deviation. Standard errors are generated using block bootstrapping following Kapetanios (2008).

A.2 Alternative specifications of baseline model

A.2.1 Global Factor

As outlined in Section 4.1, Table 4.1 reports results where we re-estimate our baseline model separately with the FCI (replacing equity volatility), with the global factor of Miranda-Agrippino and Rey (2015) (replacing equity volatility), and with all variables in annual space.

The global factor proposed in Miranda-Agrippino and Rey (2015) is extracted from a large panel of risky asset prices across various geographical areas, and is available from 1980 to 2018.[26] It uses a dynamic factor model to summarise fluctuations in global financial markets and includes asset prices traded on all the major global markets covering North and Latin America, Europe, Asia and Australia.

TABLE 4.1: Estimated impact on 5$^{\text{th}}$ percentile of GDP growth after 12 quarters

	Baseline (1)	2	3	4
Credit-to-GDP (3yr pp change)	-0.32	-0.30	-0.35	-0.27
	(-0.21, -0.43)	(-0.15, -0.46)	(-0.23, -0.46)	(0.01, -0.54)
Real House Prices (3yr growth)	-0.25	0.03	-0.15	-0.17
	(-0.04, -0.45)	(0.21, -0.16)	(0.03, -0.33)	(0.22, -0.56)
Current account (% of GDP)	-0.52	-0.63	-0.68	-0.52
	(-0.4, -0.64)	(-0.46, -0.8)	(-0.57, -0.8)	(-0.32, -0.72)
Volatility (SDs from Mean)	0.11			0.02
	(0.24, -0.01)			(0.22, -0.19)
FCI		0.09		
		(0.25, -0.08)		
Global Factor			-0.67	
			(-0.42, -0.92)	
Capital Ratio (quarterly)	0.31	0.57	0.14	
	(0.51, 0.11)	(0.73, 0.4)	(0.31, -0.03)	
Capital Ratio (annual)				0.31
				(0.58, 0.04)

Note: This table shows estimates of the average annual impact of one-standard-deviation increases in each variable on the 5$^{\text{th}}$ percentile of GDP growth over the following 12 quarters. Four separate specifications are used: (1) our baseline, (2) our baseline with the FCI replacing equity volatility, (3) our baseline with a global factor (see Miranda-Agrippino and Rey, 2015) replacing equity volatility, and (4) our baseline, but with all variables in annual space. Numbers in brackets refer to one-standard-deviation confidence bands.

[26] We thank the authors for providing us with extended data on the global factor. The time series used in Miranda-Agrippino and Rey (2015) covers the shorter period of 1990-2012.

A.2.2 Households and corporate credit

In Figure 4.1, we plot local projections showing the impact of a one-standard-deviation increase in each indicator on GDP-at-Risk in our baseline model. Figure 4.8 repeats this estimation, but splits total credit into its household and corporate credit components. The top row of Figure 4.8 presents the impact of a change in household or corporate credit-to-GDP on GDP-at-Risk, and shows that after 20 quarters, the impact of an increase in household credit on tail risk is twice as large as the impact of corporate credit. The main messages from other indicators in relatively similar to our baseline results in Figure 4.1, although the coefficient on the current account is generally smaller.

A.2.3 Single indicator models

As a cross-check on the baseline results in Figure 4.2, Figure 4.9 reports results from an alternative specification of quantile regressions where the impact of vulnerability indicators is estimated individually.[27] We obtain broadly similar results to our baseline model in this exercise. The medium-term coefficients for house price growth and the current account change very little, but the magnitude of the coefficient on credit growth increases by two-thirds.

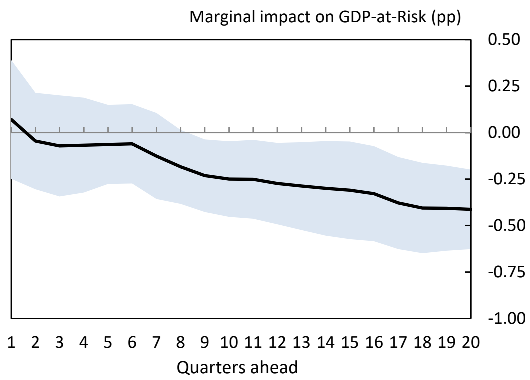
A.3 Comparing actual GDP outturns with the full predicted GDP growth distribution

Here we compare actual GDP realisations against the full predicted GDP growth distribution based on our baseline model. The 3-year-ahead forecast horizon is used such that the actual GDP growth outturn is allocated to a percentile bucket based on the GDP growth distribution predicted 3 years previously. For example, suppose that in 1994:Q3, our baseline model had predicted that the 60$^{\text{th}}$ percentile for GDP growth over the next 3 years in Country X would average 2.73% and the 70$^{\text{th}}$ percentile would average 2.88%. Then if the actual outturn for GDP growth in that country between 1994:Q3 and 1997:Q3 averaged 2.79%, then the 1997:Q3 growth observation would be allocated to the 60-70$^{\text{th}}$ percentile bucket. This process is repeated for each GDP growth observation for each country in our sample. Figure 4.10 shows that the proportion of GDP observations across all countries in our sample falling within each percentile bucket is broadly in line with the expected proportions. While this is an in-sample exercise with the coefficients coming from

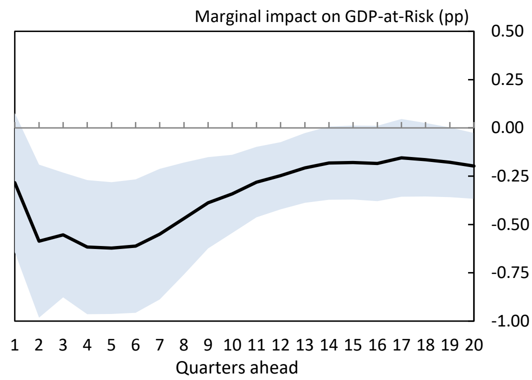
[27] These regressions with individual vulnerability indicators also include macroeconomic controls.

Figure 4.8: Baseline results with credit split into household and corporate contributions

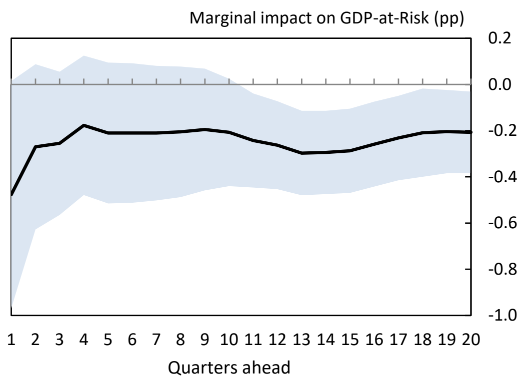
(a) Household credit-to-GDP (3 year pp change)



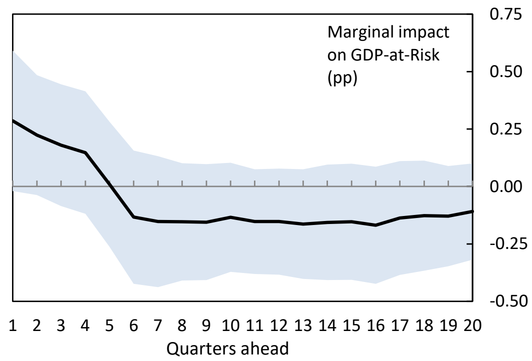
(b) Corporate credit-to-GDP (3 year pp change)



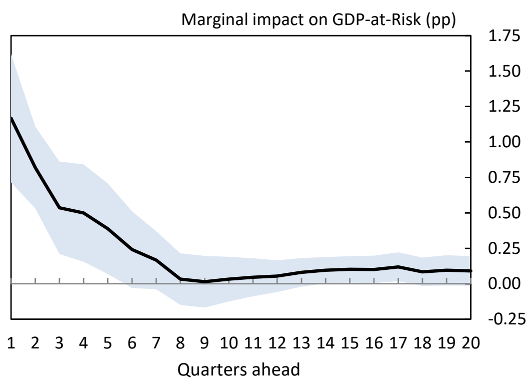
(c) Current account deficit



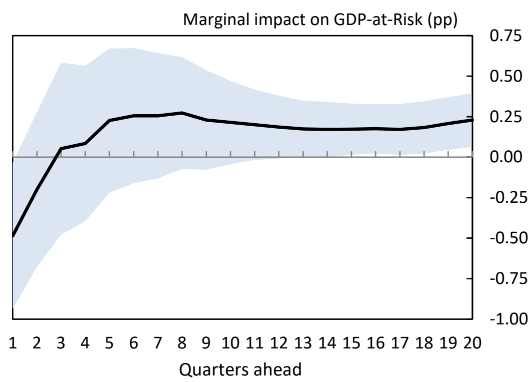
(d) Real house price growth (3 year)



(e) Volatility



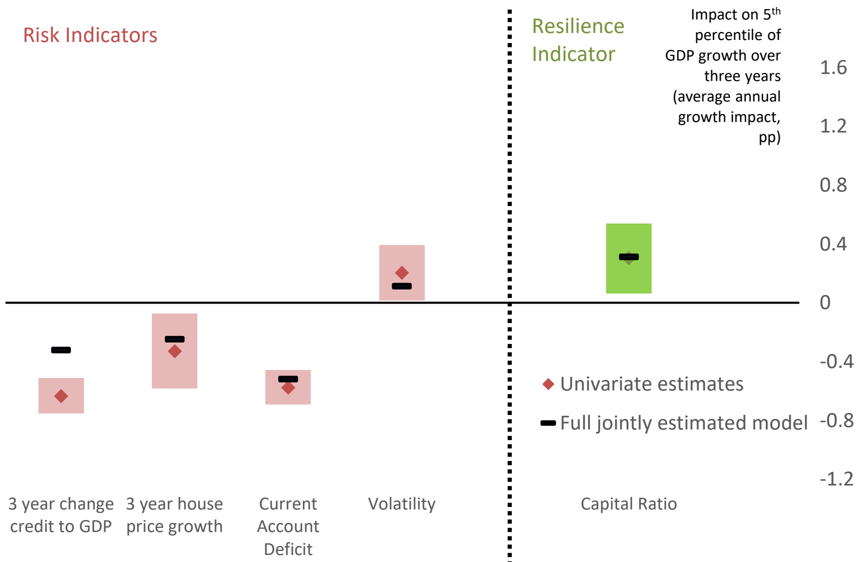
(f) Capital ratio



Note: these charts show the impact of a change in the indicator at time t on the 5^{th} percentile of real GDP growth at each horizon on the x-axis. GDP growth is measured as the average annual growth rate at each horizon. Confidence intervals represent ± 1 standard deviation. Standard errors are generated using block bootstrapping following Kapetanios (2008).

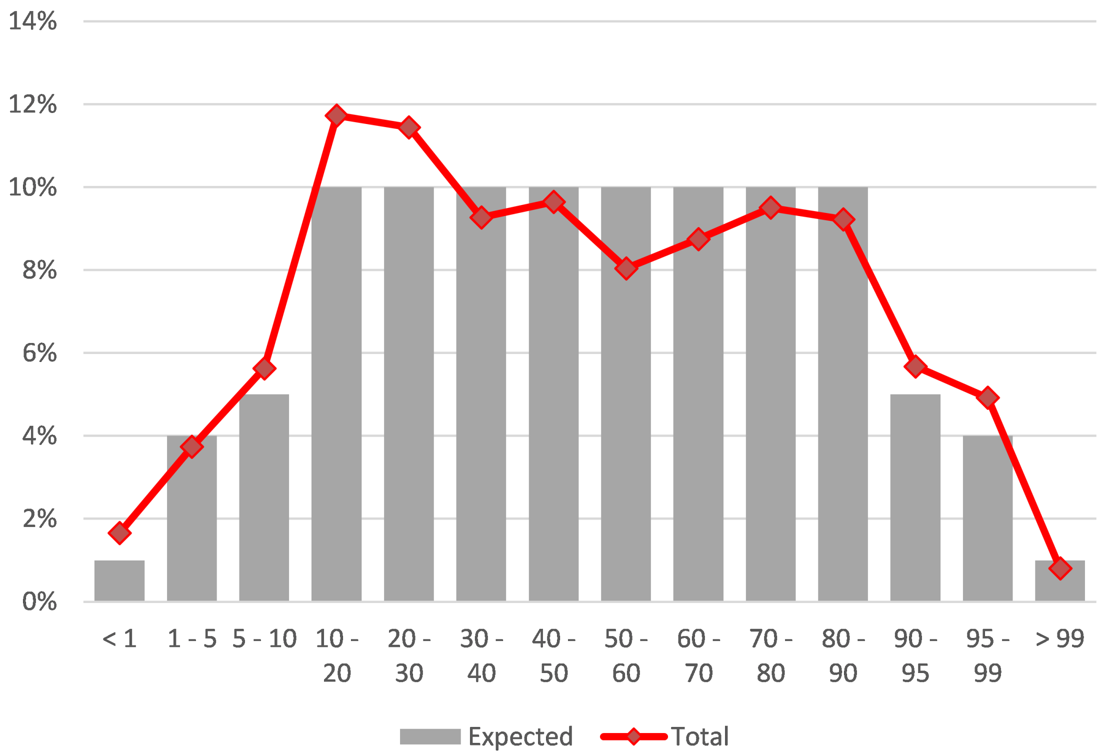
our baseline model estimated at each quantile using the full data sample, it nevertheless provides a reassuring check of the overall goodness of fit of our model.

Figure 4.9: Baseline results and single-indicator model



Note: this figure shows the impact of a one-standard-deviation increase in a given indicator at time t on the 5^{th} percentile of real GDP growth after 12 quarters. GDP growth is measured as the average annual growth rate at the 3-year horizon. Confidence intervals represent ± 1 standard deviation and correspond to the single-indicator model. Standard errors are generated using block bootstrapping following Kapetanios (2008). The coefficients labelled single indicator estimates are those obtained when each vulnerability indicator is included individually in the specification, alongside our macroeconomic controls (lagged GDP growth, inflation, and the annual change in central bank policy rate). The black bars denote the coefficients obtained from our full baseline model, where all five vulnerabilities indicators are included jointly (the results from Figure 4.2).

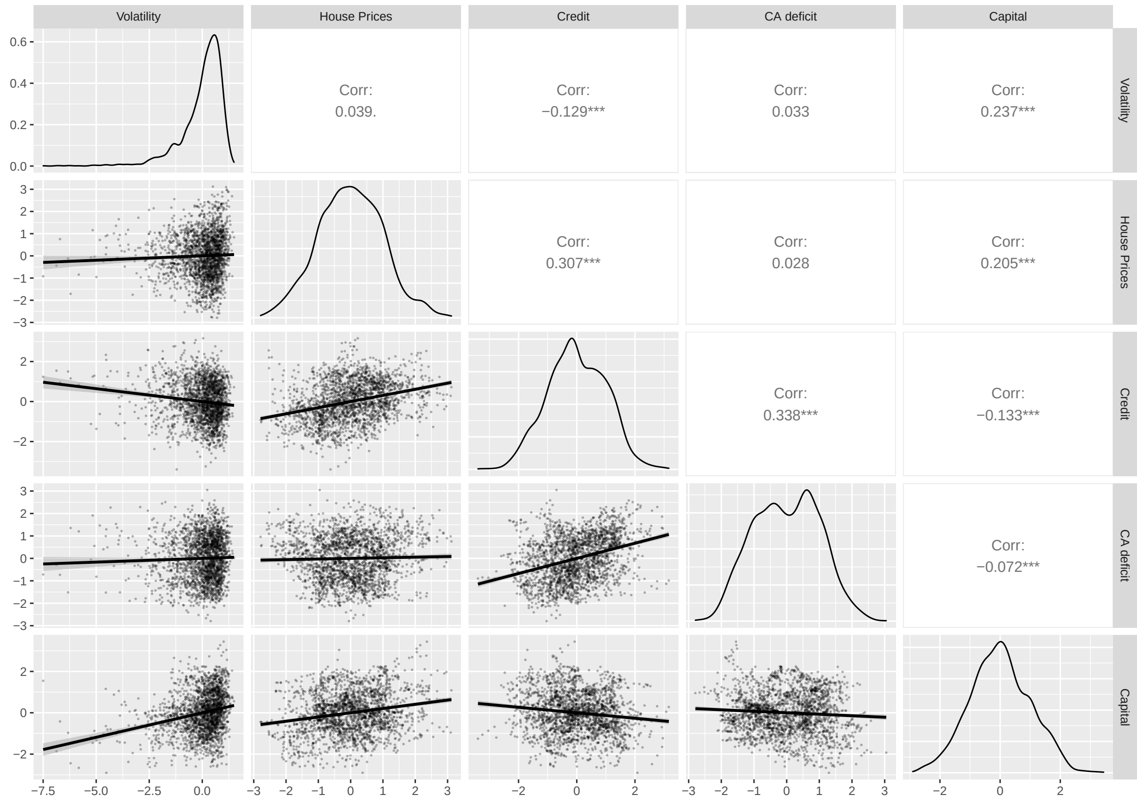
FIGURE 4.10: Proportion of actual GDP growth outturns across all countries falling into each part of the GDP distribution predicted 3 years previously



Note: the red line shows the proportion of actual GDP growth outturns falling into each percentile bucket, based on the predicted GDP distribution from 3 years earlier. The predicted GDP distribution is based on our baseline model, estimated over the full sample of countries and the full time series. The grey bars simply show the expected proportion of observations falling into each bucket (i.e. 10% to fall into each decile). Ireland's observations are excluded from the red line given a heavy loading at the extreme right-hand tail of the distribution. This reflects GDP data reclassifications and does not affect our analysis in this paper, which is focused on the left tail.

A.4 Additional charts

FIGURE 4.11: Correlation between the variables used in the quantile regressions



Note: The matrix shows the cross-correlations of each series with the others. The lower triangular part shows the scatter plots of variables with linear regression lines. Diagonal charts provide the distribution of the variables and the upper triangular reports the corresponding correlation coefficient with *s to indicate the significance level of the correlation. All variables are standardised so that they are in the form in which they enter the quantile regressions.

B Data Appendix

B.1 Capital ratios

We construct an annual cross-country measure of the tangible common equity (TCE) ratio that builds on Brooke et al. (2015). First, for each country, we obtain annual data on total assets, equity and intangible assets for each banking group operating in a given year from Thomson Reuters Worldscope. Measures of tangible assets and tangible equity for each bank are then obtained by subtracting intangible assets from each of total assets and total equity.

To account for the entry and exit of banks at different points in time within the financial system, we adopt a "chain-weighting" approach to produce a "spliced" country-level measure of tangible assets and tangible equity. For the year 2005, our spliced measure of tangible assets is simply the raw sum of tangible assets across banks in 2005 as we use 2005 as the base year. For the year 2004, the spliced measure of tangible assets is calculated as:

$$\text{Spliced TA in 04} = \text{Spliced TA in 05} \times \frac{\text{Raw 04 sum for banks operating in both 04 \& 05}}{\text{Raw 05 sum for banks operating in both 04 \& 05}}$$

Similarly for the year 2003, the formula becomes:

$$\text{Spliced TA in 03} = \text{Spliced TA in 04} \times \frac{\text{Raw 03 sum for banks operating in both 03 \& 04}}{\text{Raw 04 sum for banks operating in both 03 \& 04}}$$

The process continues back to the initial year. For years after 2005, the calculation is very similar. For example, for the year 2006:

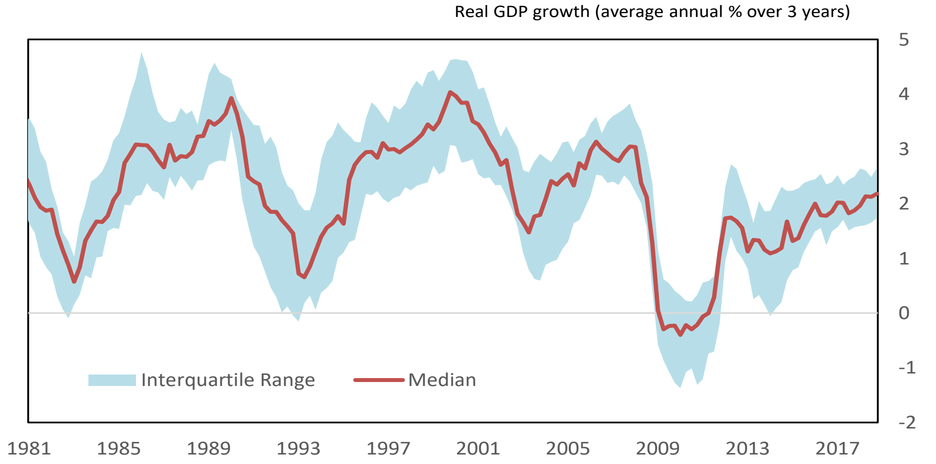
$$\text{Spliced TA in 06} = \text{Spliced TA in 05} \times \frac{\text{Raw 06 sum for banks operating in both 05 \& 06}}{\text{Raw 05 sum for banks operating in both 05 \& 06}}$$

The same construction applies for tangible equity. The TCE ratio is then computed as spliced tangible assets divided by spliced tangible equity. We apply linear interpolation to obtain quarterly values from the annual series.

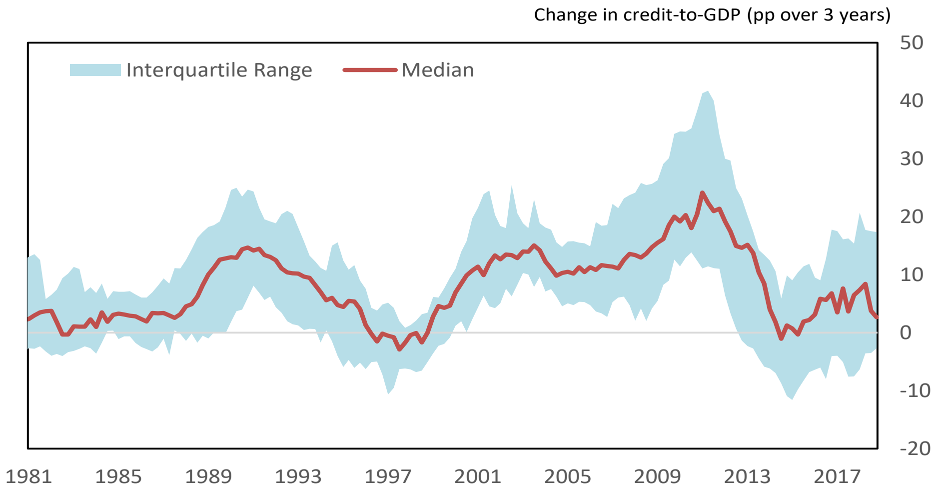
Table 4.2 documents data sources for each variable, Table 4.3 reports summary statistics on our dataset, and Figure 4.12 plots the median and interquartile range of real GDP growth, changes in credit-to-GDP, and the TCE ratio across our panel of countries. Table 4.4 reports summary statistics on the banks used to construct the capital ratios across countries, in particular summary statistics on the number of banks, market capitalisation, bank tangible assets, and total assets across the banking sector. The average number of banks per year across country-year pairs is 18, although Table 4.4 shows that there is heterogeneity across countries and over time. The US has the most banks per year with 88.6 banks on average, while Ireland has the least with 3.4 banks on average. Summary statistics at the bank level on tangible assets (in terms of local currency) and market capitalisation (in terms of US dollars for publicly-traded banks in our sample) are also reported. In addition, we report summary statistics on aggregate assets across all banks in a given country and year. For example, at end-2017, total assets in our data were £5.6 trillion in the UK, which covered 90% of total banking system assets as measured by the denominator in the Financial Policy Committee's leverage indicator.

Figure 4.12: Median and Interquartile range of selected indicators across sample of countries

(a) Real GDP growth



(b) 3-year change in credit-to-GDP



(c) Capital ratio

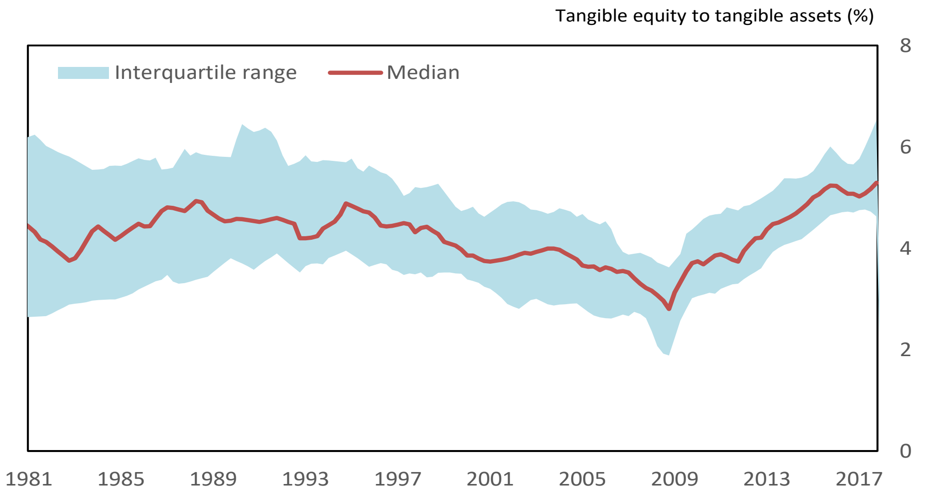


Table 4.2: Data sources

Variable	Data Source	Frequency	Notes
Real GDP	OECD	Quarterly	
Credit-to-GDP	BIS	Quarterly	3 year change in ratio of private non-financial credit to GDP
House prices	OECD	Quarterly	3 year growth in real house prices
Current Account	OECD	Quarterly	Per cent of GDP
Volatility	Datastream	Daily	Quarterly SD of daily return in national equity market
Capital Ratio	Worldscope	Annual	Ratio of tangible common equity to tangible assets
Inflation	OECD	Quarterly	Annual growth of CPI
Policy Rate	BIS	Quarterly	Annual change in central bank policy rate

Table 4.3: Summary statistics by country

		N	Mean	Std Dev.	Min	p25	p75	Max
Australia	Credit-to-GDP (3yr change)	149	10.2	10.9	-13.4	4.4	18.3	29.8
	Real House Prices (3yr growth)	149	11.5	13.6	-10.7	0.6	19.5	53.8
	Current account (% of GDP)	149	-4.3	1.2	-6.9	-5.1	-3.3	-2.1
	Volatility (SDs from Mean)	149	0.0	1.0	-7.5	-0.3	0.6	1.3
	Capital Ratio	149	5.0	0.7	3.5	4.4	5.7	6.3
	Inflation	149	4.0	3.0	-0.4	1.9	6.1	12.4
	Policy Rate (1yr change)	149	-0.4	2.9	-15.0	-1.3	0.5	7.8
Belgium	Credit-to-GDP (3yr change)	149	10.5	11.9	-12.1	2.2	16.7	47.6
	Real House Prices (3yr growth)	149	5.6	15.6	-37.9	0.4	15.3	28.7
	Current account (% of GDP)	149	1.9	2.2	-3.2	0.2	3.5	5.2
	Volatility (SDs from Mean)	149	0.0	1.0	-4.5	-0.4	0.7	1.1
	Capital Ratio	149	3.3	0.7	1.2	2.8	3.7	4.5
	Inflation	149	2.7	2.1	-1.1	1.3	3.1	9.9
	Policy Rate (1yr change)	149	-0.3	1.3	-5.0	-1.3	0.3	3.0
Canada	Credit-to-GDP (3yr change)	149	7.6	10.0	-14.5	0.5	14.9	30.6
	Real House Prices (3yr growth)	149	8.5	15.2	-25.5	-1.0	19.0	56.0
	Current account (% of GDP)	149	-1.5	2.1	-4.2	-3.3	0.5	3.0
	Volatility (SDs from Mean)	149	0.0	1.0	-6.7	-0.3	0.6	1.1
	Capital Ratio	149	3.6	0.4	2.6	3.3	3.9	4.3
	Inflation	149	3.1	2.6	-0.9	1.5	4.0	12.8
	Policy Rate (1yr change)	149	-0.3	2.1	-7.2	-1.3	0.8	8.4
Denmark	Credit-to-GDP (3yr change)	149	9.4	16.6	-13.9	-4.9	20.2	47.8
	Real House Prices (3yr growth)	149	5.4	23.5	-48.5	-14.6	21.9	57.6
	Current account (% of GDP)	149	1.8	3.7	-5.3	-1.1	3.5	9.3
	Volatility (SDs from Mean)	149	0.0	1.0	-6.9	-0.3	0.6	1.4
	Capital Ratio	149	5.4	1.4	2.8	4.3	6.6	7.9
	Inflation	149	3.0	2.5	0.2	1.7	3.4	12.2
	Policy Rate (1yr change)	149	-0.3	1.3	-6.3	-0.9	0.2	3.5
Finland	Credit-to-GDP (3yr change)	149	7.9	15.7	-45.1	3.7	15.5	48.1
	Real House Prices (3yr growth)	149	8.5	21.8	-46.7	-0.7	21.6	70.9
	Current account (% of GDP)	149	0.8	3.7	-5.8	-1.8	4.0	8.4
	Volatility (SDs from Mean)	149	0.0	1.0	-3.9	-0.4	0.7	1.2
	Capital Ratio	149	5.1	1.1	2.5	4.1	5.8	7.6
	Inflation	149	3.2	2.9	-0.5	1.2	3.9	13.8
	Policy Rate (1yr change)	149	-0.2	1.0	-4.0	-0.5	0.0	2.0

Summary statistics by country

		N	Mean	Std Dev.	Min	p25	p75	Max
France	Credit-to-GDP (3yr change)	149	7.2	6.4	-7.1	2.0	12.4	18.8
	Real House Prices (3yr growth)	149	6.0	16.4	-22.6	-7.7	20.1	44.5
	Current account (% of GDP)	149	0.0	1.3	-4.0	-0.8	0.8	3.8
	Volatility (SDs from Mean)	149	0.0	1.0	-5.1	-0.4	0.6	1.3
	Capital Ratio	149	2.8	0.7	1.4	2.5	3.2	4.1
	Inflation	149	3.0	3.1	-0.4	1.4	3.2	14.2
	Policy Rate (1yr change)	149	-0.3	1.4	-3.3	-1.2	0.2	5.6
Germany	Credit-to-GDP (3yr change)	149	1.2	6.3	-10.7	-3.1	6.6	11.7
	Real House Prices (3yr growth)	149	-0.7	6.8	-12.6	-5.9	3.7	15.2
	Current account (% of GDP)	149	2.7	3.4	-2.2	-0.9	5.7	9.1
	Volatility (SDs from Mean)	149	0.0	1.0	-5.0	-0.5	0.7	1.3
	Capital Ratio	149	2.7	0.7	1.7	2.3	2.8	5.2
	Inflation	149	2.0	1.5	-1.1	1.1	2.7	7.2
	Policy Rate (1yr change)	149	-0.2	1.1	-3.5	-0.5	0.5	2.5
Ireland	Credit-to-GDP (3yr change)	149	19.4	32.9	-43.1	-0.3	28.2	111.4
	Real House Prices (3yr growth)	149	10.4	28.4	-42.0	-10.1	29.8	73.7
	Current account (% of GDP)	149	-1.5	3.8	-12.5	-3.7	1.0	8.2
	Volatility (SDs from Mean)	149	0.0	1.0	-5.2	-0.4	0.7	1.1
	Capital Ratio	149	5.5	1.6	3.2	4.5	6.4	9.7
	Inflation	149	3.6	4.6	-2.8	1.5	4.0	23.3
	Policy Rate (1yr change)	149	-0.4	1.8	-6.8	-1.3	0.3	4.5
Italy	Credit-to-GDP (3yr change)	149	4.6	8.4	-11.7	-2.7	10.6	22.1
	Real House Prices (3yr growth)	149	4.2	24.4	-41.0	-14.5	20.5	66.7
	Current account (% of GDP)	149	-0.3	1.8	-3.7	-1.6	1.3	3.3
	Volatility (SDs from Mean)	149	0.0	1.0	-4.1	-0.5	0.7	1.5
	Capital Ratio	149	4.7	0.7	3.4	4.3	5.0	6.7
	Inflation	149	4.6	4.4	-0.3	2.0	5.5	19.6
	Policy Rate (1yr change)	149	-0.4	1.5	-6.5	-1.0	0.3	4.0
Netherlands	Credit-to-GDP (3yr change)	149	14.1	9.3	-9.9	7.2	19.4	41.0
	Real House Prices (3yr growth)	149	5.2	22.1	-48.1	-8.0	17.8	47.7
	Current account (% of GDP)	149	4.8	2.7	-0.4	2.7	6.9	10.8
	Volatility (SDs from Mean)	149	0.0	1.0	-5.2	-0.3	0.7	1.1
	Capital Ratio	149	3.8	0.8	2.5	3.0	4.5	5.5
	Inflation	149	2.1	1.6	-1.2	1.3	2.7	7.3
	Policy Rate (1yr change)	149	-0.3	1.2	-5.0	-0.8	0.3	3.0

Summary statistics by country

		N	Mean	Std Dev.	Min	p25	p75	Max
Norway	Credit-to-GDP (3yr change)	149	9.5	16.2	-22.5	-1.7	23.4	44.3
	Real House Prices (3yr growth)	149	12.1	20.6	-31.6	-0.1	26.8	68.8
	Current account (% of GDP)	149	6.8	6.0	-6.6	2.9	12.1	17.3
	Volatility (SDs from Mean)	149	0.0	1.0	-6.3	-0.3	0.6	1.3
	Capital Ratio	149	4.5	1.1	1.6	3.9	5.4	6.8
	Inflation	149	3.7	3.1	-1.4	1.9	4.5	14.7
	Policy Rate (1yr change)	149	-0.2	1.8	-6.0	-0.8	0.3	5.5
Spain	Credit-to-GDP (3yr change)	149	7.6	21.1	-35.3	-3.3	23.7	53.8
	Real House Prices (3yr growth)	149	11.5	33.8	-43.5	-13.5	34.1	111.7
	Current account (% of GDP)	149	-2.4	3.0	-10.2	-3.9	-0.5	2.3
	Volatility (SDs from Mean)	149	0.0	1.0	-4.2	-0.4	0.7	1.5
	Capital Ratio	149	5.1	0.8	3.1	4.7	5.5	6.9
	Inflation	149	4.6	3.8	-1.1	2.3	6.1	16.1
	Policy Rate (1yr change)	149	-0.4	2.8	-10.6	-1.5	0.5	11.7
Sweden	Credit-to-GDP (3yr change)	149	10.4	16.7	-26.6	0.4	16.7	63.1
	Real House Prices (3yr growth)	149	8.8	21.7	-34.2	-7.0	27.1	42.9
	Current account (% of GDP)	149	2.7	3.4	-3.1	-0.2	5.4	8.4
	Volatility (SDs from Mean)	149	0.0	1.0	-4.5	-0.5	0.7	1.2
	Capital Ratio	149	3.6	0.7	1.8	3.2	3.9	5.0
	Inflation	149	3.3	3.5	-1.2	0.8	5.2	14.8
	Policy Rate (1yr change)	149	-0.3	3.9	-32.0	-1.0	1.0	30.0
Switzerland	Credit-to-GDP (3yr change)	149	7.1	9.1	-9.7	0.2	13.4	30.0
	Real House Prices (3yr growth)	149	3.8	12.9	-26.1	-3.2	11.9	35.0
	Current account (% of GDP)	149	7.8	3.7	-0.6	4.5	10.9	15.1
	Volatility (SDs from Mean)	149	0.0	1.0	-4.8	-0.3	0.6	1.3
	Capital Ratio	149	4.4	1.8	1.7	2.9	6.3	7.0
	Inflation	149	1.7	2.0	-1.4	0.4	2.8	7.1
	Policy Rate (1yr change)	149	-0.1	1.1	-2.4	-0.9	0.3	3.0
UK	Credit-to-GDP (3yr change)	149	7.0	11.7	-20.2	-0.2	16.5	23.4
	Real House Prices (3yr growth)	149	13.4	23.0	-28.2	-6.0	31.1	69.4
	Current account (% of GDP)	149	-2.1	1.7	-5.9	-3.5	-0.7	2.3
	Volatility (SDs from Mean)	149	0.0	1.0	-5.8	-0.3	0.7	1.2
	Capital Ratio	149	4.1	0.9	1.8	3.5	4.7	5.5
	Inflation	149	3.4	2.6	0.0	1.6	4.4	15.2
	Policy Rate (1yr change)	149	-0.4	1.8	-5.0	-1.3	0.5	4.9
USA	Credit-to-GDP (3yr change)	149	4.1	8.8	-18.2	-1.0	11.6	18.4
	Real House Prices (3yr growth)	149	2.7	11.9	-22.3	-5.9	13.6	22.0
	Current account (% of GDP)	149	-2.6	1.5	-6.1	-3.3	-1.6	0.3
	Volatility (SDs from Mean)	149	0.0	1.0	-6.2	-0.2	0.6	1.2
	Capital Ratio	149	5.5	1.1	2.8	4.8	6.0	8.1
	Inflation	149	3.1	2.0	-1.6	1.9	3.7	12.5
	Policy Rate (1yr change)	149	-0.4	2.0	-8.9	-1.3	0.8	8.2
All Sample	Credit-to-GDP (3yr change)	2384	8.6	15.2	-45.1	-0.2	15.4	111.4
	Real House Prices (3yr growth)	2384	7.3	20.9	-48.5	-5.7	19.5	111.7
	Current account (% of GDP)	2384	0.9	4.5	-12.5	-2.3	3.5	17.3
	Volatility (SDs from Mean)	2384	0.0	1.0	-7.5	-0.3	0.7	1.5
	Capital Ratio	2384	4.3	1.4	1.2	3.3	5.1	9.7
	Inflation	2384	3.2	3.1	-2.8	1.4	3.9	23.3
	Policy Rate (1yr change)	2384	-0.3	2.0	-32.0	-1.0	0.5	30.0

Table 4.4: Banking system data: summary statistics by country

		N	Mean	Std Dev.	Min	p25	p75	Max
Australia	Number of banks per year	38	9.3	3.1	4	8	12	13
	Market capitalisation ($m)	308	15449	25048	8.2	419	17133	123289
	Tangible assets (bn. AUD)	353	150.5	243.4	0.08	4.9	148.6	965.4
	Aggregate total assets (bn. AUD)	38	1414	1303	48.0	378	2564	3913
Belgium	Number of banks per year	38	4.0	2.1	2	2	6	8
	Market capitalisation ($m)	148	5740	9197	21.4	238	6386	47703
	Tangible assets (€bn)	153	113.5	144.9	0.008	5.7	212.7	644.8
	Aggregate total assets (€bn)	38	459	301	66.3	160	695	1000
Canada	Number of banks per year	38	10.0	2.2	6	9	12	14
	Market capitalisation ($m)	333	15089	23892	12.8	1015	16252	113668
	Tangible assets (bn. CAD)	380	178.0	259.1	0.0001	6.5	253.1	1317
	Aggregate total assets (bn. CAD)	38	1798	1474	312	558	2718	5368
Denmark	Number of banks per year	38	27.3	13.4	4	20	38	44
	Market capitalisation ($m)	1009	583	2870	2.5	19.4	159	34810
	Tangible assets (bn. DKK)	1039	69.4	380.2	0.3	1.4	11.4	3532
	Aggregate total assets (bn. DKK)	38	1904	1417	80.6	845	3606	4089
Finland	Number of banks per year	38	3.9	1.3	2	3	5	6
	Market capitalisation ($m)	122	820	1145	6.7	161	961	6417
	Tangible assets (€bn)	147	11.2	13.0	0.04	1.7	14.8	62.2
	Aggregate total assets (€bn)	38	43.6	23.5	10.5	18.7	63.9	77.6
France	Number of banks per year	38	23.1	9.4	7	18	30	42
	Market capitalisation ($m)	745	4991	13710	6.3	215	1915	98706
	Tangible assets (€bn)	879	120.6	329.9	0.07	3.1	46.5	2059
	Aggregate total assets (€bn)	38	2807	2085	265	1118	5518	6012
Germany	Number of banks per year	38	17.3	6.9	8	11	25	29
	Market capitalisation ($m)	530	5213	10171	2.3	360	4228	66666
	Tangible assets (€bn)	659	118.9	270.5	0.003	8.4	108.5	2184
	Aggregate total assets (€bn)	38	2071	1176	360	842	3009	4020
Ireland	Number of banks per year	38	3.4	1.3	1	3	4	6
	Market capitalisation ($m)	93	3603	4701	1.7	392	4759	20628
	Tangible assets (€bn)	131	50.8	54.6	0.1	7.0	80.4	196.4
	Aggregate total assets (€bn)	38	176	162	5.7	38.0	281	554
Italy	Number of banks per year	38	26.7	10.0	9	19	35	43
	Market capitalisation ($m)	904	3896	9686	0.1	339	3061	110084
	Tangible assets (€bn)	1015	48.1	122.8	0.004	3.4	38.0	1009
	Aggregate total assets (€bn)	38	1301	878	93.6	420	2264	2459
Netherlands	Number of banks per year	38	7.3	2.7	2	6	10	11
	Market capitalisation ($m)	121	9205	16446	35.4	232	9301	99754
	Tangible assets (€bn)	279	162.9	277.9	0.2	6.0	141.4	1311
	Aggregate total assets (€bn)	38	1198	984	142	365	1733	3451

Banking system data: summary statistics by country

		N	Mean	Std Dev.	Min	p25	p75	Max
Norway	Number of banks per year	38	17.4	8.0	4	13	23	29
	Market capitalisation ($m)	613	767	3144	3.8	20.0	253	30175
	Tangible assets (bn. NOK)	660	85.8	300.2	0.2	5.1	48.5	2692
	Aggregate total assets (bn. NOK)	38	1494	1359	63.8	559	2791	4316
Spain	Number of banks per year	38	14.4	5.6	6	9	19	23
	Market capitalisation ($m)	501	8239	19225	5.3	424	6115	136121
	Tangible assets (€bn)	546	84.6	197.7	0.04	2.9	62.6	1392
	Aggregate total assets (€bn)	38	1232	1216	46.3	182	2385	3386
Sweden	Number of banks per year	38	4.4	1.2	3	4	5	7
	Market capitalisation ($m)	136	12873	12886	22.2	2528	20170	54071
	Tangible assets (bn. SEK)	168	1217	1453	1.8	140.3	2037	6368
	Aggregate total assets (bn. SEK)	38	5413	4958	229	1000	11265	13886
Switzerland	Number of banks per year	38	19.4	6.4	4	20	23	26
	Market capitalisation ($m)	585	5926	15539	6.1	140	2282	117800
	Tangible assets (bn. CHF)	738	92.1	281.8	0.9	5.3	25.2	2378
	Aggregate total assets (bn. CHF)	38	1800	1130	233	621	2589	3954
UK	Number of banks per year	38	12.0	1.9	8	11	13	15
	Market capitalisation ($m)	343	26813	42647	4.0	1784	40748	210836
	Tangible assets (£bn)	456	226.5	399.1	0.003	23.6	206.9	2375
	Aggregate total assets (£bn)	38	2741	2487	123	645	5621	8186
USA	Number of banks per year	38	88.6	44.3	38	45	132	162
	Market capitalisation ($m)	3308	7905	28444	0.001	240	3599	366302
	Tangible assets ($bn)	3365	61.8	236.1	0.003	3.7	31.7	2517
	Aggregate total assets ($bn)	38	5616	3686	1041	2427	9810	12111

Note: This table provides summary statistics across countries on the banks used to construct the capital ratio series in Sections 2 and B.1. "Number of banks per year" shows summary statistics on the number of annual bank observations available for a given country. "Market capitalisation" shows summary statistics on the market capitalisation at the bank level for those banks in our sample that are publicly traded, and is expressed in terms of US dollars. "Tangible assets" shows summary statistics on total tangible assets at the bank level, where tangible assets are calculated as total assets minus intangible assets and are expressed in terms of the local currency. "Aggregate total assets" gives the sum of total assets across the banks in a given country and year, and is expressed in terms of the local currency.