

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

# Essays in Public Economics and Development

Milad Khatib-Shahidi

A thesis submitted to the Department of Economics of the  
London School of Economics for the degree of Doctor of  
Philosophy, London, March 2018

# Declaration

I certify that the thesis I have presented for examination for the MPhil/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).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

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

The following disclaimer applies to chapter 1 of the thesis which uses DWP data: "This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates."

The following disclaimer applies to chapter 2 of the thesis which uses HMRC data: "This work contains statistical data from HMRC which is Crown Copyright. The research datasets used may not exactly reproduce HMRC aggregates. The use of HMRC statistical data in this work does not imply the endorsement of HMRC in relation to the interpretation or analysis of the information."

# Abstract

I present three essays in this thesis. The first essay provides novel empirical evidence on the evolution of the incentive cost of unemployment benefits during an unemployment spell. Theoretical arguments have been proposed for both inclining and declining benefit profiles. However, empirical evidence on how the incentive cost of unemployment benefits may vary over the spell, which is a key input in evaluating the time profile of benefits, is limited and mixed. I estimate the incentive cost of benefits paid at various points during an unemployment spell and find that the elasticity of unemployment duration with respect to benefits and the incentive cost of UI are smaller for benefits paid later in the spell. I argue that the decline in incentive costs is driven by partially myopic job-search behaviour and non-stationarities in the dynamics of job search. The second essay provides quasi-experimental evidence of the short-term and long-term effects of fiscal stimulus programs in the UK housing and auto markets. In an influential work Mian and Sufi (2012) argue that such temporary incentives are ineffective in boosting market activity in the long-term. I show that a temporary tax cut in UK housing market has had considerable long-term effects. I argue, using a dynamic search model with frictions, that the magnitude of the long-term effect of a stimulus is directly related to its duration. The third essay shows that frequent repayment can act as a screening device in micro-lending under individual liability. A tight repayment schedule can be used to screen out "risky" borrowers. Borrowers with more volatile profits would prefer contracts with higher interest rate but more flexible repayment schedule, while "safe" borrowers can afford to repay more frequently. I show that frequent repayment can be used to design a menu of contracts that achieves a separating equilibrium.

# Acknowledgements

I am extremely grateful to my supervisors Camille Landais and Maitreesh Ghatak for all their support and encouragement throughout my PhD.

I have benefited from helpful discussions with Johannes Spinnewijn, Henrik Kleven, Arthur Seibold, Hiu Fung Tam, Panos Mavrokonstantis, Michel Dummar Azulai and Mohammad Vesal. I am also indebted to the HMRC Datalab and ADRN staff for providing access to HMRC and DWP datasets used in chapters 1 and 2.

I would like to express my eternal gratitude to my parents for their unconditional love and support. I would also like to thank my sister, Maryam, whose kindness and caring knows no bounds.

Finally, I cannot begin to express my gratitude to my beloved wife, So-mayeh, for her patience, sacrifice and encouragements through the ups and downs of these years. She never stopped believing in me even when I stopped believing in myself.

# Contents

<b>1</b>	<b>Dynamic Incentive Costs of Unemployment Insurance: Evidence From the UK</b>	<b>9</b>
1.1	Introduction . . . . .	9
1.2	Conceptual Framework: Moral Hazard Cost of UI . . . . .	13
1.3	Institutional Background and Data . . . . .	15
1.4	Empirical Strategy . . . . .	18
1.5	Results . . . . .	20
1.5.1	Graphical Evidence . . . . .	20
1.5.2	Duration Responses and Moral Hazard Costs . . . . .	22
1.5.3	Implications for the Optimal Time Profile of Benefits . . . . .	24
1.6	Why Does Moral Hazard Cost Decline? . . . . .	25
1.6.1	Are Job Seekers Forward-Looking? . . . . .	26
1.6.2	Non-stationarity . . . . .	29
1.7	Identification and Robustness Checks . . . . .	31
1.8	Conclusion . . . . .	32
1.9	Figures and Tables . . . . .	34
<b>2</b>	<b>Effects of Fiscal Stimulus: Evidence From the UK</b>	<b>53</b>
2.1	Introduction . . . . .	53
2.2	The Effects of UK Stamp Duty Reform . . . . .	55
2.2.1	Context and Data . . . . .	55
2.2.2	Empirical Strategy . . . . .	56

2.2.3	Results . . . . .	58
2.3	Model . . . . .	59
2.3.1	Setup . . . . .	60
2.3.2	Extensive Margin Response . . . . .	61
2.3.3	Search Effort . . . . .	62
2.4	Total Response And The Share of Extensive Margin . . . . .	63
2.5	Conclusion . . . . .	64
2.6	Figures and Tables . . . . .	66
<b>3</b>	<b>Frequent Repayments as a Screening Device in Micro-lending</b>	
	<b>Under Individual Liability</b>	<b>78</b>
3.1	Introduction . . . . .	78
3.2	One-Shot Repayment . . . . .	80
3.2.1	The Environment . . . . .	80
3.2.2	The One-Period Model . . . . .	81
3.3	Frequent Repayment as a Screening Device . . . . .	82
3.3.1	Returns of Investment Projects . . . . .	82
3.3.2	Preferences over Contracts and the Single-Crossing Property . . . . .	84
3.3.3	Optimal Contracting with Gradual Repayment . . . . .	86
3.4	Conclusion . . . . .	89
3.5	Figures . . . . .	90
	<b>References</b>	<b>93</b>
	<b>A Proofs for Chapter 3</b>	<b>97</b>

# List of Figures

1.1	Variation in Time Profile of UI Benefits . . . . .	34
1.2	Average Censored Durations of Unemployment . . . . .	35
1.3	Average Duration of Unemployment . . . . .	36
1.4	Elasticity of Duration w.r.t Benefits Paid at Different Spell Lengths. . . . .	37
1.4	Elasticity of Duration w.r.t Benefits Paid at Different Spell Lengths (continued) . . . . .	38
1.5	Moral Hazard Cost of UI Benefits Paid at Different Spell Lengths	39
1.5	Moral Hazard Cost of UI Benefits Paid at Different Spell Lengths (continued) . . . . .	40
1.6	Controlling for the Effect of Age on Unemployment Duration . .	41
1.7	Survival Curves for Cohorts Starting Spells 12 and 24 Weeks Before 25 . . . . .	42
1.8	Survival Curves for Cohorts Starting Spells 8, 16 and 24 Weeks Before Turning 25 . . . . .	43
1.9	Effect of change in current and future benefits on hazards out of unemployment. . . . .	44
1.10	Effect of changes in current benefits on exit rates from unem- ployment. . . . .	45
1.11	Exit Rate From Unemployment . . . . .	46
1.12	Number of UI Claims in Weekly Bins of Age . . . . .	47
1.13	Duration of Unemployment Around Birthdays Other Than 25 .	48

2.1	Property transactions tax liability . . . . .	68
2.2	Number of transactions in the treated and control price brackets.	69
2.3	Number of transactions in the treated and control price brackets.	70
2.4	The difference in activity between the treated and the control price brackets. . . . .	71
2.5	Cumulative difference in market activity between treated and control price brackets. . . . .	72
2.6	The value of entering the market over time. . . . .	73
2.7	Search effort over time. . . . .	74
2.8	Simulated number of transactions over time. . . . .	75
2.9	Number of matches and the number of entries to the market, in excess of high-tax steady state. . . . .	76
2.10	Share of extensive margin response in the total short-term re- sponse. . . . .	77
3.1	Indifference curves of risky and safe borrowers . . . . .	90
3.2	The constraints of the contracting problem . . . . .	91
3.3	Optimal separating contracts . . . . .	92



# List of Tables

1.1	Duration Elasticities and Moral Hazard Costs . . . . .	49
1.2	Duration Elasticities With Respect to Current and Future Benefits	50
1.3	Robustness to the Method of Estimating the Counterfactual . .	51
1.4	Placebo Elasticities Estimated Around Birthdays Other Than 25	52
2.1	Property transactions tax rates . . . . .	66
2.2	Difference-in-difference regression results . . . . .	67

# Chapter 1

## Dynamic Incentive Costs of Unemployment Insurance: Evidence From the UK

### **1.1 Introduction**

The aim of unemployment insurance (UI) is to provide insurance against job loss while maintaining the incentives for workers to search for a new job. As these two forces act in opposing directions, the trade-off between them is the key to the optimal design of UI policy. As time profile of UI benefits affects the dynamics of household behaviour, including consumption and job search, the design of optimal UI policy requires knowledge of the dynamics of incentive cost and insurance value of UI over the unemployment spell. Although these are ultimately empirical questions, we have almost no empirical evidence on how these costs and benefits evolve over the unemployment spell. Consequently, little is known about the optimal path of UI benefits. This is evident from the considerable variation in the time profile of benefits across countries which reflects the lack of consensus on whether benefits should increase or decrease over time.

This paper provides novel evidence on how the incentive costs of UI change over the unemployment spell, using unique quasi-experimental variation in the

time profile of UI benefits in the UK. This variation is due to age-dependence of UI benefits and provides exceptionally rich variation in benefit profiles across cohorts. I exploit the fact that UI claimants receive a lower amount of weekly benefits so long as they are under 25 and the level of benefits increases automatically on the week in which the claimant turns 25. This creates a step-shaped benefit profile, in which the location of the step will depend on the age of claimant at the start of the spell. To estimate the effects of variation in benefits at different unemployment durations, I compare the outcomes for different cohorts who start their spells at different ages before 25 to those who start just after turning 25 (who will face a flat benefit profile).

I exploit the flexible yet simple nature of this variation to estimate the incentive cost of UI benefits paid at different unemployment durations and provide a rich image of how these costs evolve over the spell.

Seminal papers in the theoretical UI literature have studied the optimal time profile of benefits. However, the results from these studies are model-dependent and cannot easily be connected to the data (Shavell and Weiss 1979; Hopenhayn and Nicolini 1997; Werning 2002). Recently, a growing empirical literature has employed the sufficient statistics approach to evaluate social insurance policies based on high-level and easily estimable statistics. However, this literature has focused almost exclusively on policies with a constant benefit level and has been mostly silent about how the incentive cost and insurance value of UI might change with the duration of unemployment.

One important exception is Kolsrud et al. (2018) who derive sufficient statistics for characterizing the optimal time profile of benefits and implement their framework in the context of the Swedish UI system. Although this provides the first, and to my knowledge, the only attempt at revealing the dynamics of incentive costs of UI, Kolsrud et al. (2018) are constrained by the nature of the variation they can use in Swedish UI policy. In particular, they are only able to estimate the incentive cost associated with UI benefits paid

in the first 20 weeks, and those paid thereafter. In contrast, the rich variation in the UK benefit profile used in this paper lends itself to a much more flexible estimation strategy and yields a richer image of the evolution of incentive costs.

In this paper, I estimate the incentive cost of UI benefits associated with benefits paid at different parts of the unemployment spell. I use administrative data from the UK and exploit the variation in UI benefit profiles created by the dependence of UI benefit levels on the age of claimants. I first estimate the elasticity of total duration of unemployment with respect to benefits paid over different parts of the unemployment spell. These elasticities measure the magnitude of the behavioural response to UI benefits paid at different points of a spell. Similar to the standard Bailey-Chetty formula, the incentive cost of benefits paid at time  $t$  of the unemployment spell is fully captured by the corresponding fiscal externality, that is, the effect of increasing those benefits on government budget. I calculate these incentive costs for benefits paid in each part of the spell based on the corresponding estimates of duration elasticities.

To provide a more detailed account of how the incentive cost changes with duration of unemployment, and to ensure robustness of the findings, I exploit the flexibility of the policy variation and hypothetically divide the benefit profile into periods of various lengths (e.g., 8-weeks, 12-weeks, 3 months, etc.) and repeat the estimation for each configuration.

I find that the elasticity of unemployment duration with respect to benefits paid at time  $t$  during the spell consistently falls with  $t$ , across all configurations. In other words, UI benefits paid later in the spell induce a weaker behavioural response. This, in turn, implies that the incentive cost associated with benefits paid at longer durations is smaller than benefits paid earlier in the spell. This has important implications for the optimal time profile of UI benefits. Both theory and existing empirical evidence suggest that the insurance value of UI increases with duration of unemployment, as agents run down their as-

sets. The combination of decreasing incentive costs and increasing insurance values would imply that increasing the tilt of the benefit profile, i.e. providing relatively more generous long-term benefits, would increase welfare (Kolsrud et al. 2018).

The declining pattern of incentive costs and the implied inclining benefit profile is in contrast with the finding in the theoretical UI literature that incentive costs rise over the spell and the UI policy must, therefore, be less generous towards the long-term unemployed to incentivize job search (Shavell and Weiss 1979; Hopenhayn and Nicolini 1997). This theoretical result is derived in a model with forward-looking agents and in a stationary environment. To find the reason for the discrepancy between the theory and my results, I empirically investigate both of these assumptions for job seekers in the UK. I show that individuals fail to act in a forward-looking manner and do not respond to changes in future benefits in advance. I also find that job finding rates become less responsive to benefits at longer durations, which points towards non-stationarities in the job-search environment. Incorporating such non-stationarities in a dynamic job search model and assuming non-forward-looking search<sup>1</sup>, indeed results in rising incentive costs (Kolsrud et al. 2018). Therefore, these findings can help explain why, in spite of theory, incentive costs fall over the unemployment spell.

This paper contributes to three literatures. First, the sufficient statistics approach to design and evaluation of UI has focused almost exclusively on overall costs and benefits of constant benefit profiles. I contribute to this literature by revealing how the incentive cost of UI evolves over the unemployment spell. Second, I contribute to the literature on the effects of UI on labour supply (Rothstein 2011; Lalive, Van Ours, and Zweimüller 2006; Lalive 2008) by providing estimates of the response of unemployment duration to UI benefits at different durations. Finally, this paper contributes to the lit-

---

<sup>1</sup>Note that failure to respond to future benefits does not necessarily imply myopia, as it is also consistent with lack of knowledge about the benefit schedule.

erature on the effects of behavioural factors (such as inattention and present bias) on individuals' responses and how this may alter the design of optimal policy (Chetty, Looney, and Kroft 2009; Chetty, Friedman, and Saez 2013) by providing evidence of myopic job-search.

The rest of the paper is organized as follows. Section 1.2 will briefly lay out the conceptual framework for identifying incentive costs of UI. Section 1.3 describes the data and the institutional background of unemployment insurance in the UK. Section 1.4 discusses the empirical strategy while section 1.5 presents the main results on duration elasticities and moral hazard costs. Section 1.6 provides evidence of the mechanisms behind the findings. Section 1.8 concludes.

## **1.2 Conceptual Framework: Moral Hazard Cost of UI**

This section briefly lays out the conceptual framework and derives the incentive cost of UI. The derivation presented here closely follows Kolsrud et al. (2018).

The general insight developed by Baily (1978) and Chetty (2006) can also be applied to the case of a dynamic UI policy. Suppose the UI policy consists of  $n$  parts and consider the effect on social welfare of increasing the benefit in part  $k$ ,  $b_k$ , by  $db_k$ . Such an increase affects social welfare through three channels. First, there is the direct effect of benefits on the utility of the unemployed which increases social welfare proportional to their marginal utility of consumption. Second, the government needs to raise an additional tax revenue of  $db_k D_k$ ,  $D_k$  being the expected time spent by the unemployed in part  $k$ , to cover the mechanical cost of increasing benefits, in the absence of any behavioural responses. However, individuals will respond to this more generous benefit profile by staying unemployed for longer. This creates the third effect which is the incentive cost of UI benefits. Due to this behavioural response in unemployment durations, the government needs to raise the taxes

further to finance the cost of paying UI benefits for longer<sup>2</sup>.

More specifically, as the argument above shows, the welfare cost of increasing  $b_k$  is captured by its effect on government's budget. Let  $T$  denote the lifetime of agents,  $D_k$  denote the expected time spent by the agent claiming  $b_k$ , and  $D = \sum_{k=1}^n D_k$  denote the expected duration of total unemployment. Given UI policy  $\{b_1, \dots, b_n\}$ , government's budget can be written as:

$$G = (T - D)\tau - D_1 b_1 - D_2 b_2 - \dots - D_n b_n \quad (1.1)$$

Denoting by  $\epsilon_{D_l, b_k}$  the elasticity of expected duration in part  $l$  with respect to benefits in part  $k$ , the welfare cost of increasing  $b_k$  is therefore:

$$\begin{aligned} \frac{\partial G}{\partial b_k} = & -D_k \times \left\{ 1 + \sum_{l=1}^n \frac{D_l (b_l + \tau)}{D_k b_k} \epsilon_{D_l, b_k} \right\} \\ & -D_k \times \{1 + MH_k\} \end{aligned} \quad (1.2)$$

This means that when increasing benefits  $b_k$  by one unit, the required rise in taxes is  $MH_k$  times more than the implied mechanical cost. Equation (1.2) shows that the mechanical cost, captured by the first term, is proportional to the expected duration spent claiming  $b_k$ . The moral hazard (incentive) cost  $MH_k$  depends on how the expected unemployment duration in each part of the spell responds to the change in  $b_k$ .

Although equation (1.2) suggests that to calculate  $MH_k$  one would, in principle, need to estimate  $\epsilon_{D_l, b_k}$  for all combinations of  $l$  and  $k$ , starting from a flat benefit profile ( $b_k = \bar{b}$ ) the moral hazard cost of an increase in  $b_k$  will simplify to

---

<sup>2</sup>There is, in principle, a fourth effect as higher  $b_k$  induces longer unemployment and this can affect the utility of the unemployed. However, as Chetty (2006) shows, envelope conditions of individuals' optimization problem imply that this only has a second order effect on their welfare as they have already optimized over benefits. More generally, job-seekers may respond to changes in benefits along various margins. However, as long as they take these variables into account in their optimization problem, the readjustments along these margins in response to changes in benefits will only have a second order effect on their welfare.

$$MH_k = \epsilon_{D,b_k} \cdot \frac{D}{D_k} \cdot \frac{\bar{b} + \tau}{\bar{b}} \quad (1.3)$$

which only depends on the elasticity of total duration of unemployment. Since UI benefits are constant for the majority of claimants in the UK, I will use Equation (1.3) in the empirical analysis. Equation (1.3) defines the disincentive cost of increasing  $b_k$  as the ratio of the behavioural cost ( $D\epsilon_{D,b_k}$ ) to the mechanical cost ( $D_k$ ), as proposed by Schmieder and Wachter (2017). Finally, the ratio  $(\bar{b} + \tau)/\bar{b}$  corrects this for the lost tax revenue from longer unemployment spells.

The evolution of moral hazard costs has important implications for the design of optimal benefit profile. A decreasing sequence of moral hazard costs would imply that it is less costly to increase long-term benefits than short-term ones. Everything else equal, this would indicate that introducing an inclining benefit profile that offers relatively more generous benefits to the long-term unemployed, would improve welfare.

Kolsrud et al. (2018) show that in a stationary environment and with forward-looking job-seekers, the moral hazard cost  $MH_k$  is increasing in  $k$ . However, as they also note, this result crucially depends on stationarity of the environment as well as forward-looking job-search behaviour. Therefore, presence of non-stationarities, prevalent in many countries (Machin and Manning 1999) or deviations from forward-looking search due to behavioural factors (e.g. present bias) can overturn this result. Therefore, how moral hazard costs evolve over the spell is eventually an empirical question.

### 1.3 Institutional Background and Data

The unemployment insurance in the United Kingdom is known as Jobseeker's Allowance (JSA). There are two types of JSA: income-based and contribution-based. I will focus on income-based claimants for reasons explained below<sup>3</sup>.

---

<sup>3</sup>Contribution-based JSA is not means-tested and is available to those who have paid sufficient National Insurance contributions in the two years prior to job loss. The amount



Income-based JSA is a means-tested benefit, available to claimants with a low level of assets<sup>4</sup> and can be claimed indefinitely, provided the claimants meet the job-search requirements of the program.

The weekly amount of benefits does not depend on prior earnings and is determined solely based on the personal circumstances of claimants. The benefits are made up of two elements: a personal allowance plus certain premiums (disability, dependent children etc.). I will focus on income-based JSA claimants who only qualify for the personal allowance component.

The key variation I will use is due to the fact that the amount of benefit (the personal allowance component) depends on the age of individuals. In particular, claimants over 25 are paid at a higher rate than those under-25<sup>5</sup>. Notice that what determines the amount of benefit is not the age at the time of making a JSA claim. Rather, the amount of benefit would increase automatically if claimants turn 25 while on unemployment benefits. This creates a step-shaped time profile for benefits and the exact location of this step depends on how long before turning 25 a job-seeker starts her spell. This creates rich variation in the time profile of benefits across cohorts of claimants which I will exploit for my empirical strategy.

I use administrative data covering the universe of JSA claims in Great Britain<sup>6</sup> between January 2001 and April 2016. This dataset is an extract of the National Benefits Database, held by the Department for Work and Pension. The data includes information on start and end dates of claims, birthday of the claimant, type of claim (contribution- or income-based), amount of benefit paid, and reason for end of claim.

---

of contribution-based benefit is the same as that of income-based. Its duration however, is limited to 182 days, unlike the income-based which can be claimed indefinitely.

<sup>4</sup>Less than £16,000 in savings.

<sup>5</sup>For example, the weekly benefits in 2015 for 18-24 and over-25 claimants were £57.90 and £73.10, respectively. The amount of benefit is updated every year. However, since JSA benefits are closely indexed with inflation, the real rate of benefit almost stays constant over the period of this study.

<sup>6</sup>This includes England, Wales and Scotland, but not Northern Ireland.

**New Deal for Young People.** In order to assist young job seekers back into employment, the UK government introduced the New Deal for Young People (NDYP) in 1998 which was in place until 2011. To separate the effects of this program from the effects of changes in benefits, for some of the empirical results I will only use post-2011 spells. The rest of this section provides further details on this.

Under NDYP, job seekers under 25 who had been unemployed for 6 months, and those older than 25 who had been unemployed for 18 months, would receive job search assistance as well as training. This would affect individuals' duration of unemployment. In particular, job seekers who started a spell more than 6 months before turning 25 would have reached 6 months of unemployment before turning 25 and would enter NDYP, while people who started their spell less than 6 months before turning 25 would already be older than 25 in 6 months and would not enter NDYP (unless they stayed unemployed for 18 months). Because of this, for spells that start before 2011, I will only be able to use the variation in benefit profiles within the first 6 months of the spell, i.e. claimants who start their spell within 6 months of turning 25. The reason is that claimants who start their spells more than 6 months before turning 25 were treated differently and are not comparable to those starting within 6 months of turning 25.

To make sure that NDYP does not confound the estimates, I will present two sets of results for duration responses and moral hazard costs. First, I will focus on the response to UI benefits in the short-term using spells that start within 6 months of turning 25 from all years. The second set of results will estimate the response to benefits for longer unemployment durations, but will only use spells after 2011 to avoid the effects of NDYP.

## 1.4 Empirical Strategy

The incentive cost of UI depends crucially on how the duration of unemployment responds to changes in UI benefits. Therefore, as Equation 1.3 also shows,  $\epsilon_{D,b_k}$  is the key statistic we need to estimate to evaluate the moral hazard cost of UI in part  $k$  of the spell. This section discusses the empirical strategy for estimating  $\epsilon_{D,b_k}$  as well as other components of  $MH_k$ . Given that the current benefit profile in the UK is flat (that is, for claimants aged 25 and above), the following empirical strategy is designed to estimate the duration responses and moral hazard costs around the current flat benefit profile. This is to ensure that the resulting estimates reflect the costs of reforming the existing flat benefit profile in the UK.

The variation in the time profile of UI benefits is due to the fact that unemployed individuals under 25 are paid a lower rate of benefit than those over 25. In particular, if a claimant turns 25 during an unemployment spell, the benefits will automatically increase during the week in which she turns 25.

Panel A of Figure 1.1 shows the level of benefit as a function of age around 25-th birthday. Consider the cohort of claimants who make a UI claim  $k$  periods before turning 25. These are marked as cohort  $k$  in panel A of Figure 1.1. Panel B shows the resulting step-shaped benefit profile that cohort  $k$  face. Now consider job-seekers who start their spell immediately after turning 25 I will refer to these as cohort 0, as their spell starts 0 periods before 25. Cohort 0 will face a flat benefit profile at the higher rate as shown by the blue section in panel C. Panel D shows the flat benefit profile of cohort 0 along with that of cohort  $k$ . The two cohorts will face different levels of UI benefits for the first  $k$  periods, but same benefits thereafter. Denoting the benefits paid during the first  $k$  periods by  $\bar{b}_k$ , this creates variation in  $\bar{b}_k$  which can be used to estimate  $\epsilon_{D,\bar{b}_k}$ , the elasticity of unemployment duration with respect to benefits in the first  $k$  periods. In particular, I will exploit this variation in

regressions of the following form, on spells from cohorts 0 and  $k$ <sup>7</sup>:

$$\ln(D_{i,t}) = \beta_0 + \delta_k \ln(\bar{b}_{k,t}) + \theta_t + u_{i,t} \quad (1.4)$$

where  $D_{i,t}$  is the duration of unemployment for spell  $i$  from year  $t$ ,  $\bar{b}_{k,t}$  is the benefits of the first  $k$  periods in year  $t$  and  $\theta_t$  are year fixed effects.

The resulting  $\hat{\delta}_k$  coefficients provide estimates of  $\epsilon_{D,\bar{b}_k}$ , the elasticity of duration with respect to all benefits up to period  $k$ . To recover the duration elasticities with respect to  $b_k$  only, notice that when evaluated at a flat benefit profile

$$\epsilon_{D,b_k} = \epsilon_{D,\bar{b}_k} - \epsilon_{D,\bar{b}_{k-1}} \quad (1.5)$$

which provides a simple method for recovering  $\epsilon_{D,b_k}$ <sup>8</sup>.

Equation (1.4) yields consistent estimates if  $\bar{b}_k$  are exogenous with respect to the error term, which requires the identifying assumption that, in the absence of the increase in benefits at 25, age at the start of spell would not be correlated with duration of unemployment. Therefore, a possible threat to identification is the case where age directly affects duration of unemployment. If, for example, unemployment duration tends to increase with age regardless of UI benefits, estimates of duration elasticities from equation (1.4) will be biased upwards. This is because part of the difference in unemployment durations across cohorts would be driven by differences in age and not UI benefit.

To address this, I report a second set of estimates where I adjust the observed durations for the effect of age as follows. I allow for age to be related to duration of unemployment, but assume that, in the absence of the increase in benefits at 25, expected duration would evolve smoothly with age at start

---

<sup>7</sup>To increase statistical power, I use a bandwidth of 4 weeks and include spells that start within 2 weeks on either side of cohort  $k$  in that cohort. Cohort 0 in the regression consists of spells starting up to 10 weeks after 25. Results are robust to changing these bandwidths.

<sup>8</sup>An alternative strategy would be to compare cohort  $k$  to cohort  $k-1$  and directly estimate  $\epsilon_{D,b_k}$ , as these cohorts face the same benefit profile in all time periods except for period  $k$ . However, both of these cohorts are paid the lower rate for the first  $k-1$  periods and the higher rate after period  $k$ . Therefore the resulting estimate with respect to  $b_k$  would not correspond to variations around a flat benefit profile.

of spell. Based on this identifying assumption, I estimate the counterfactual relationship between age and expected duration by fitting a polynomial to the average duration of unemployment excluding spells that start up to 3 years before 25 as depicted in Figure 1.6<sup>9</sup>. This counterfactual predicts how the expected duration of unemployment would evolve with age if all cohorts faced the same flat benefit profile at the higher (25+) rate. Therefore, the difference between the predicted durations for cohort 0 and cohort  $k$  measures how much of the observed difference in durations can be attributed to difference in age. So, I subtract this from observed durations to remove the effect of age and then use this corrected unemployment duration as the dependent variable in a regression similar to equation (1.4) to estimate duration elasticities.

To estimate  $MH_k$ , we also need estimates of expected duration of total unemployment,  $D$ , and expected time spent in part  $k$ ,  $D_k$ . Since the estimation strategy is designed to provide estimates around a flat benefit profile, I estimate  $D$  by the average duration for spells that start just after 25, and  $D_k$  by the average time spent claiming  $b_k$ . I use  $\frac{b_+}{b} = 1.06$  which corresponds to the tax to benefit ratio that would balance government's budget during the period covered by the sample.

For empirical implementation, given the flexible nature of the available policy variation, I examine different configurations by dividing the unemployment spell into periods of various lengths and report results for each configuration to ensure the robustness of the results.

## 1.5 Results

### 1.5.1 Graphical Evidence

Figure 1.2 provides graphical evidence on the response of unemployment duration to variation in UI benefits at different durations. The figure shows average

---

<sup>9</sup>A more conservative method would be to only use spells that start after 25 and estimate a linear counterfactual. This yields similar results which I report in section 1.7.

duration of unemployment, censored at 6 months<sup>10</sup>, in weekly bins of age at the start of spell. It clearly illustrates some interesting points. First, average unemployment duration is stable for spells starting sufficiently earlier than 25 (say, before 23) and those starting immediately after 25, and does not seem to vary across cohorts. Also, as we approach the age of 25 and get closer to higher benefits, the duration of unemployment increases steadily, as expected.

More importantly, the rate of increase in average unemployment duration steadily rises as we approach 25, forming a convex trend. This curvature of unemployment durations graph reveals an important point about how UI benefits at different durations affect duration of unemployment. To see this, consider two consecutive cohorts that start their spells  $n$  and  $n - 1$  weeks before turning 25. These cohorts will face the same benefit profile except for week  $n$ , during which cohort  $n - 1$  have already turned 25 and receive benefit at the higher (25+) rate while cohort  $n$  are still paid at the lower rate. The difference in unemployment durations of these cohorts would reflect the effect of a change in UI benefits during week  $n$  of the spell. The fact that durations rise faster for cohorts closer to 25 implies that this effect is stronger for smaller values of  $n$ . That is, unemployment durations respond more strongly to benefits paid earlier in the spell.

Given the flexible nature of the variation in benefit profiles, I will report estimates of duration responses and moral hazard costs that correspond to periods of various lengths. In accordance with this, Figure 1.3 shows the uncensored average duration of unemployment in 8 and 12 week steps. Each point in these graphs corresponds to spells that start in the same week of age, as before, but adjacent points are now 8 and 12 weeks apart in panels A and B respectively.

Note that the large increase in average duration that happens 6 months before 25, which was removed through censoring in Figure 1.2 is due to the

---

<sup>10</sup>Durations are censored at 6 months to remove the effect of the NDYP program on durations. Uncensored durations exhibit similar patterns, as discussed below.

NDYP program and is not related to the variation in benefit profile. This will be accounted for in the empirical strategy, as discussed in section 1.4. Disregarding this sharp rise, average unemployment duration exhibits the same convex trend as individuals approach 25. Figure 1.3 also confirms that the mean unemployment duration does not vary systematically with age for spells starting after 25 or sufficiently earlier than 25. I will nevertheless, also report estimates that adjust for the possible effect of age on durations.

### 1.5.2 Duration Responses and Moral Hazard Costs

Figure 1.4 and Table 1.1 present estimates of duration elasticities with respect to benefits paid at different parts of the spell, both with and without adjustment for the effect of age of unemployment duration. Panels A and B of Figure 1.4 show duration responses to benefits paid within the first 6 months of the spell, for 8 and 12-week periods. These estimates reveal that even in the short term, the elasticity of unemployment duration with respect to UI benefits consistently declines over the spell. For 8-week periods, the elasticity of duration (before adjusting for age) with respect to benefits falls from 0.387 (0.023) for the first period to 0.103 (0.027) in the second and 0.026 (0.025) for the third period. Elasticities with respect to 12-week periods also fall from 0.433 (0.022) in the first period to 0.083 (0.026) in the second period, confirming that duration responses are indeed smaller for later benefits.

While these results focus on the first six months of the spell, it would also be interesting to see how the unemployment duration responds to changes in benefits over longer time periods. Panel C of Figure 1.4 reports similar estimates, comparing the effect of benefits in the first 6 months (26 weeks) to those of the next 18 months, using spells that start after 2011. These estimates once again show that unemployment duration is more responsive to UI benefits in the first 6 months than all benefits paid over the next 18 months.

One interesting aspect of this finding is that it can help explain why existing estimates of duration elasticity with respect to potential benefit duration

are typically somewhat smaller than elasticities with respect to benefit level (Schmieder and Von Wachter 2016). This is consistent with the above results because extension of potential benefit duration can be seen as an increase in benefits paid in later parts of the spell, as opposed to an increase in overall benefit level which changes benefits paid relatively earlier.

Based on these estimates of duration elasticities, we can calculate the moral hazard costs associated with increasing benefits for any part of the spell using equation (1.3). Table 1.1 and Figure 1.5 present estimates of moral hazard costs that correspond to the three configurations considered previously. Panels A and B of Figure 1.5 show the moral hazard cost of increasing benefits in 8 and 12 week periods within the first 6 months of the spell. The implied moral hazard cost falls over the spell across all configurations. It falls from  $MH_1 = 1.24$  (0.075) for increasing benefits paid in the first 8-weeks of spell (not adjusted for age) to  $MH_3 = 0.22$  (0.216) for the third period (i.e. weeks 17 to 24). This means that in order to increase  $b_1$  by 1 percent, the government would need to levy 1.24 times more resources from tax payers than the implied mechanical cost, while increasing  $b_3$  would only require 0.22 more in tax revenues than the mechanical cost of doing so.

As the results show, controlling for the effect of age yields slightly smaller elasticity and moral hazard cost estimates, but does not change the results qualitatively. Figure 1.6 shows that while duration of unemployment starts to increase with age around the age of 26, there does not seem to be a relationship between age and unemployment duration for spells starting up to a year after 25. To control for the effect of age, I exclude spells that start up to 3 years before 25 and fit a polynomial to the observed average durations from the remaining cohorts. This provides an estimated counterfactual which captures the relationship between age and unemployment duration. I then use this counterfactual to adjust observed spell lengths for the effect of age. Figure 1.6 shows this. Section 1.4 and Figure 1.6 provide further details.



### 1.5.3 Implications for the Optimal Time Profile of Benefits

The declining pattern in moral hazard costs has important implications for the optimal time profile of benefits. In understanding these implications it is important to keep one caveat in mind. Given that the identification strategy used here is local and based on the response of young UI claimants to benefits, it may be challenging to generalize the policy conclusions to older claimants. However, it should be noted that a substantial portion of UI claims are made by young claimants; Almost half of JSA claims in the sample are made by claimants under the age of 30. So, even if one believes that the results are only valid locally, they are still of significance as the age group studied constitute a large fraction of the unemployed and claim a considerable portion of government's total UI budget.

The moral hazard costs,  $MH_k$ , capture the loss in social welfare resulting from increasing benefits in part  $k$  of spell. To evaluate the (local) optimality of each  $b_k$  one would also need an estimate of the consumption smoothing gains to social welfare due to increasing  $b_k$ . Nevertheless, when consumption smoothing gains and moral hazard costs evolve in opposite directions, one can still draw important policy recommendations about the overall time profile of benefits. Since both theory and existing empirical evidence imply that consumption smoothing gains of UI increase over the spell<sup>11</sup> I assume that this is true in the UK as well. This will allow me to determine how the current benefit profile can be modified to increase social welfare.

In particular, consider a simple two part policy that pays benefits  $b_1$  in the first part of spell and  $b_2$  thereafter. Let  $CS_k$  denote the consumption

---

<sup>11</sup>Assuming job seekers are not liquidity-constrained and use their liquid assets to smooth consumption during the unemployment spell, the marginal utility of consumption and the value of UI benefits increases over time, as they deplete their assets. The only existing empirical estimate of consumption drop during the spell that I know of is that of Kolsrud et al. (2018) who take the consumption implementation approach (Gruber 1997) and estimate how consumption changes over the spell. They find that the consumption smoothing gain for benefits paid after the first 20 weeks is twice that of benefits paid in the first 20 weeks.

smoothing gain of increasing  $b_k$ . Kolsrud et al. (2018) show that whenever  $\frac{CS_2}{MH_2} > \frac{CS_1}{MH_1}$  an increase in  $b_2$  accompanied by a reduction in  $b_1$  improves social welfare. In this case, because consumption smoothing gains increase and the moral hazard costs decrease over the spell, this result implies that introducing an inclining benefit profile would improve social welfare.

## 1.6 Why Does Moral Hazard Cost Decline?

Influential papers in the theoretical literature on dynamic UI have examined the optimal time profile of UI benefits in stationary job-search models with forward-looking agents (Shavell and Weiss 1979; Hopenhayn and Nicolini 1997) and have found that the moral hazard cost of UI increases over the spell and therefore optimal UI benefits must be decreasing to provide the right incentives for the unemployed to search for jobs. However, this result crucially depends on the stationarity and forward-looking search assumptions which may not necessarily hold in practice.

In this section I will examine both of these assumptions empirically and show how deviations from these assumptions overturn this result and lead to decreasing moral hazard costs in the present context. I first show that job seekers do not fully respond to changes in future benefits and act in a seemingly myopic manner. This eliminates the forward-looking mechanism behind increasing moral hazard costs. I then present evidence of non-stationarities in job search environment. The combination of non-forward-looking search and non-stationary forces can explain why moral hazard costs of UI benefits decrease over the spell.

To intuitively understand the role these assumptions play in the theoretical prediction about increasing moral hazard costs, consider the following simplified version of the argument presented in Kolsrud et al. (2018). Consider the effects of changes in  $b_t$  and  $b_{t+1}$  on unemployment duration. The effect of a change in  $b_{t+1}$  can be decomposed into two components. The first compo-

ment is the effect of  $b_{t+1}$  on  $D_1$  and the second component is its effect on the remaining duration after period 1, conditional on still being unemployed after period 1. But notice that in a stationary environment, the effect of  $b_{t+1}$  on the remaining duration (after period 1) is the same as the effect of  $b_t$  on total duration<sup>12</sup>. Therefore, the effect of  $b_{t+1}$  on unemployment duration exceeds that of  $b_t$  due to its first component.

Consider now what happens if these assumptions are violated. To the extent that job-seekers exhibit myopia with respect to benefits, the first component of the effect of  $b_{t+1}$  becomes smaller. In the extreme case of full myopia, this component is zero and the moral hazard cost stays constant throughout the spell. Furthermore, suppose unemployment dynamics are non-stationary in the sense that probability of finding a job in later parts of spell becomes less responsive to changes in benefits. In that case, the second component of the effect of  $b_{t+1}$  will become smaller than the total effect of  $b_t$ . With the first component equal to zero due to myopia and the second one smaller than the effect of  $b_t$  due to non-stationarity, moral hazard cost now becomes decreasing over the spell. Therefore, determining the gradient of moral hazard costs of UI benefits is ultimately an empirical question.

### 1.6.1 Are Job Seekers Forward-Looking?

In this section I present evidence of lack of forward-looking job-search by showing that job-seekers barely respond to changes in future benefits. Note that the absence of response to future benefits does not necessarily imply myopic behaviour, as it is also consistent with lack of knowledge about the benefit path. I cannot distinguish between these two possibilities, but since I do find some limited response to changes in benefits in the near future in some settings (see Figure 1.9), lack of knowledge about benefit path does not seem to be a plausible explanation. But it must be noted that if non-forward-looking is

---

<sup>12</sup>This is because the stationarity assumption implies that the effect of a change in  $b_{t+1}$  on the continuation value of a job-seeker in period 2, is the same as the effect of a change in  $b_t$  on the continuation value at the beginning of the spell.

driven by ignorance about benefit path the welfare implications of the analysis cannot be generalized to other contexts where agents are fully aware of the benefit path.

To test the extent of forward-looking search among UK job-seekers, I start by comparing the job finding rates early in the spell of cohorts who start on the same level of benefit but face different benefit paths in the future. Assuming job seekers are forward-looking, we expect the cohort with higher future benefits to choose a relatively lower level of search effort early in the spell, in anticipation of their higher future benefits.

More specifically, consider two cohorts who start their spells 12-weeks and 24-weeks before turning 25 (the "12-week" and the "24-week" cohorts, respectively). The benefits for these two cohorts are shown in panel A of Figure 1.7. They receive the same benefits for the first 12 weeks but face different benefit profiles after that. Since the 12-week cohort will receive higher UI benefits in weeks 13-24, forward-looking search implies that their search effort should be lower from the beginning compared to the 24-week cohort. Panel B of Figure 1.7 shows the survival curves for the 12-week and the 24-week cohorts. The vertical line marks week 23 of the spell, when the 12-week cohort start receiving higher benefits. Even though these cohorts expect different levels of benefits in the future, their job finding rates are indistinguishable for the first 12-weeks. Interestingly, the two survival curves diverge immediately after week 12, when the 12-week cohort start receiving higher benefits. In other words, although job seekers react to higher current benefits by lowering their search effort, this suggests that they fail to respond to changes in future benefits.

To verify this, Figure 1.8 compares survival rates for cohorts that start their spells 8, 16 and 24 weeks before 25. Vertical lines mark the weeks in which the older cohort starts receiving higher benefits, namely weeks 8 and 16 in panels A and B respectively. Figure 1.8 confirms that the corresponding

pairs of cohorts have nearly identical job finding rates as long as they receive the same benefits and the older cohort lower their search effort only after they receive the higher benefits. Therefore, the survival curves initially coincide but start diverging once the benefit paths of the two cohorts diverge.

To provide more rigorous evidence of this seemingly myopic search behaviour, I estimate the effect of changes in benefits at different parts of the spell on hazard rates out of unemployment. The results are shown in Figure 1.9. Each panel in this figure shows the effect of changing UI benefits in an 8-week-long part of the spell on the probability of leaving unemployment in 4 week periods<sup>13</sup>. The shaded area in each panel indicates the period for which the UI benefit changes. The results in Figure 1.9 suggest that, for the most part, hazard rates out of unemployment fail to respond to changes in future benefits. More specifically, almost all hazard responses in weeks prior to the rise in benefit (that is, to the left of shaded area in each graph) are insignificant.

Additional evidence of this seemingly myopic search behaviour can be found by examining the effect of future benefits on duration of unemployment in the first part of the spell. Higher search effort early in the spell would increase chances of finding a job early on and decrease  $D_1$ <sup>14</sup>. Therefore, we would expect  $\epsilon_{D_1, b_k} > 0$  not just for  $k = 1$  but also for  $k > 1$ .

Table 1.2 presents estimates of  $\epsilon_{D_1, b_k}$ , elasticity of  $D_1$  with respect to benefits in part  $k$  of spell, for the first 6 months of the spell and for 8 and 12-week periods<sup>15</sup>. The same strategy as described in section 1.4 is used to estimate these elasticities, with  $D_1$  as the outcome variable instead of  $D$ . The positive and significant estimates in row 1 indicate that higher current benefits

---

<sup>13</sup>Results for changes in benefits over 12-week-long periods are similar.

<sup>14</sup>It is straightforward to show that duration of unemployment in part  $k$  is simply the sum of survival probabilities in each period within part  $k$ . That is,  $D_k = \sum_{t=B_{k-1}+1}^{B_k} S_t$ , with  $S_t = \prod_{j=1}^{t-1} (1 - s_j)$ , where  $S_t$  is the probability of survival until time  $t$ , and  $s_t$  is the search effort at time  $t$ .

<sup>15</sup>Once again, the estimation is restricted to the first six months to avoid the distortions caused by NDYP.

in the first part of spell indeed result in longer unemployment in that part. However, the very small and insignificant estimates of rows 2 and 3 indicate that, contrary to what we would expect from forward-looking agents, future benefits do not seem to have an effect on  $D_1$ .

The above evidence suggest that job seekers, to a large extent, fail to react to changes in future benefits. Whether this is due to truly myopic behaviour or reflects a lack of knowledge about benefit profile is difficult to determine in this context. However, regardless of the underlying mechanism, the absence of forward-looking response to future benefits eliminates the driving force behind the increase in moral hazard costs over the unemployment spell.

### 1.6.2 Non-stationarity

The second important assumption that underpins the theoretical prediction about rising moral hazard costs is that unemployment dynamics are stationary. In particular, the effect of benefits on hazard rate out of unemployment is assumed to be the same at all unemployment durations. In other words, the short-term and the long-term unemployed are assumed to be equally responsive to changes in their benefits. This may however not hold empirically if, for instance, job opportunities tend to deteriorate with spell length.

A large empirical literature has documented and studied non-stationarities in unemployment (Wolpin 1987; Blau and Robins 1986; Van den Berg 1990). Non-stationarities can arise from heterogeneity among job seekers or could be due to true duration dependence<sup>16</sup>.

It must be noted that the present characterization of moral hazard cost is robust to the nature of non-stationarities. This is because, as discussed in section 1.2, similar to the original Baily-Chetty model, the application of

---

<sup>16</sup>Heterogeneity in employability of job seekers would imply that more employable individuals exit unemployment faster and the composition of the pool of job seekers changes over time, as the share of less employable individuals increases. This change in composition lowers the overall job finding rate and results in declining exit rates from unemployment. True duration dependence can arise when job opportunities become more scarce at longer spell lengths, so that the job finding rate declines with unemployment duration.

envelope conditions of agent's optimization problem guarantees that the moral hazard cost is fully captured by the behavioural revenue effect of benefits. In particular, the moral hazard cost can be estimated based solely on the response of unemployment durations to benefits and average durations, regardless of the mechanism that drives the non-stationarities (heterogeneity or true duration dependence).

Figure 1.10 shows the effect of changes in benefits over 12-week periods on hazard rates out of unemployment, estimated for 4-week intervals. It clearly shows that the effect of benefits on job finding hazards gets weaker at longer spell lengths, meaning that job-seekers become less responsive to changes in benefits later in the spell.

Another testable implication of stationarity is that, with a flat benefit profile, exit rate from unemployment will be constant and will not depend on the duration of unemployment. The reason is that under stationarity the continuation value of the dynamic search problem will be independent of the past unemployment duration.

Figure 1.11 shows exit rates from unemployment for job seekers who start their spell immediately after turning 25. These individuals face a constant benefit profile. Yet, as documented in other contexts, Figure 1.11 shows that exit rate from unemployment in the UK decreases with duration of unemployment. This indicates that non-stationary forces are indeed in action.

These two features of UK unemployment dynamics, i.e. seemingly myopic search behaviour and the non-stationary unemployment environment, explain why moral hazard cost of UI benefits decreases over the spell. The former eliminates the force that would increase the moral hazard costs, as higher future benefits affect search in prior time periods, while the latter reduces later moral hazard costs even further as job seekers in later periods become less responsive to changes in benefits. These forces are strong enough in the present context to make moral hazard costs decrease over the unemployment

spell.

## 1.7 Identification and Robustness Checks

Identification of elasticities in equation (1.4) relies on exogeneity of individuals' cohort, i.e. age at the start of spell. One possible threat to identification, the case where age might be directly correlated with unemployment duration, was addressed in section 1.4 by explicitly accounting for the effect of age on durations. This was done by estimating a counterfactual relationship between age and duration using a flexible polynomial and spells that started before 22 or after 25. As a robustness check, Table 1.3 control for age in a more conservative way, namely, by estimating a linear counterfactual using only spells that start after 25, and reports duration elasticities and moral hazard costs. The results confirm that estimates of duration responses and moral hazard costs are robust to how the counterfactual is estimated.

Another potential source of endogeneity arises if 1) individuals are able to influence when they start a spell<sup>17</sup> (i.e. their cohort), and 2) doing so is correlated with unemployment duration. For example, if more educated under-25s who lose their jobs close to turning 25 chose to delay making a claim until after turning 25, and if education was correlated with duration of unemployment, then equation (1.4) would result in inconsistent elasticity estimates.

This section addresses this latter concern. As shown below, there are distortions in the distribution of number of claims around birthdays. However, I will argue that this cannot be driven by manipulation due to financial considerations. Furthermore, I will show that these distortions, regardless of their nature, do not affect unemployment durations. Therefore, individuals' cohort is indeed exogenous.

Figure 1.12 shows the number of UI claims in weekly bins. Vertical lines

---

<sup>17</sup>Note that workers who leave employment voluntarily do not qualify for UK Jobseekers Allowance. Therefore, it is not possible to time the start of spell by choosing when to quit.



mark the week of claimants' birthday. Number of claims drops sharply just before each birthday and rises above its long run trend immediately afterwards. At first, this may seem to suggest that those close to a birthday tend to delay claiming UI until after their birthday, presumably to receive the higher weekly benefits. However, this cannot be the case for at least two reasons. First, the fact that similar distortions happen around all birthdays indicates that the observed pattern at 25 cannot be explained by financial incentives, as there is no change in financial incentives (minimum wages, benefits, etc.) at other ages. Second, even at the age of 25, job-seekers have no incentive to delay their UI claim as income-based JSA can be claimed indefinitely and regardless of age. Therefore, it is highly unlikely for the distortion around 25 to be the result of individuals' choice. So, which cohort one ends up in is not correlated with individual characteristics<sup>18</sup>.

Nevertheless, one might still be concerned that this "birthday effect", regardless of its nature, might bias the estimates of the duration elasticities and moral hazard costs. To rule this out, I examine the average duration of unemployment around other birthdays for up to 3 years either side of 25 in Figure 1.13. Unemployment duration evolves smoothly around all other birthdays and shows no sign of being affected by this "birthday effect". I also run placebo regressions at birthdays other than 25 and test for the presence of a birthday effect. Table 1.4 presents the results of these placebo regressions.

## 1.8 Conclusion

This paper has presented novel evidence of the evolution of behavioural costs of unemployment insurance over the unemployment spell. The moral hazard costs, as measured here, are sufficient statistics for local welfare analysis and fully capture the welfare costs of providing more generous benefits at each point during the unemployment spell. The results clearly indicate that the

---

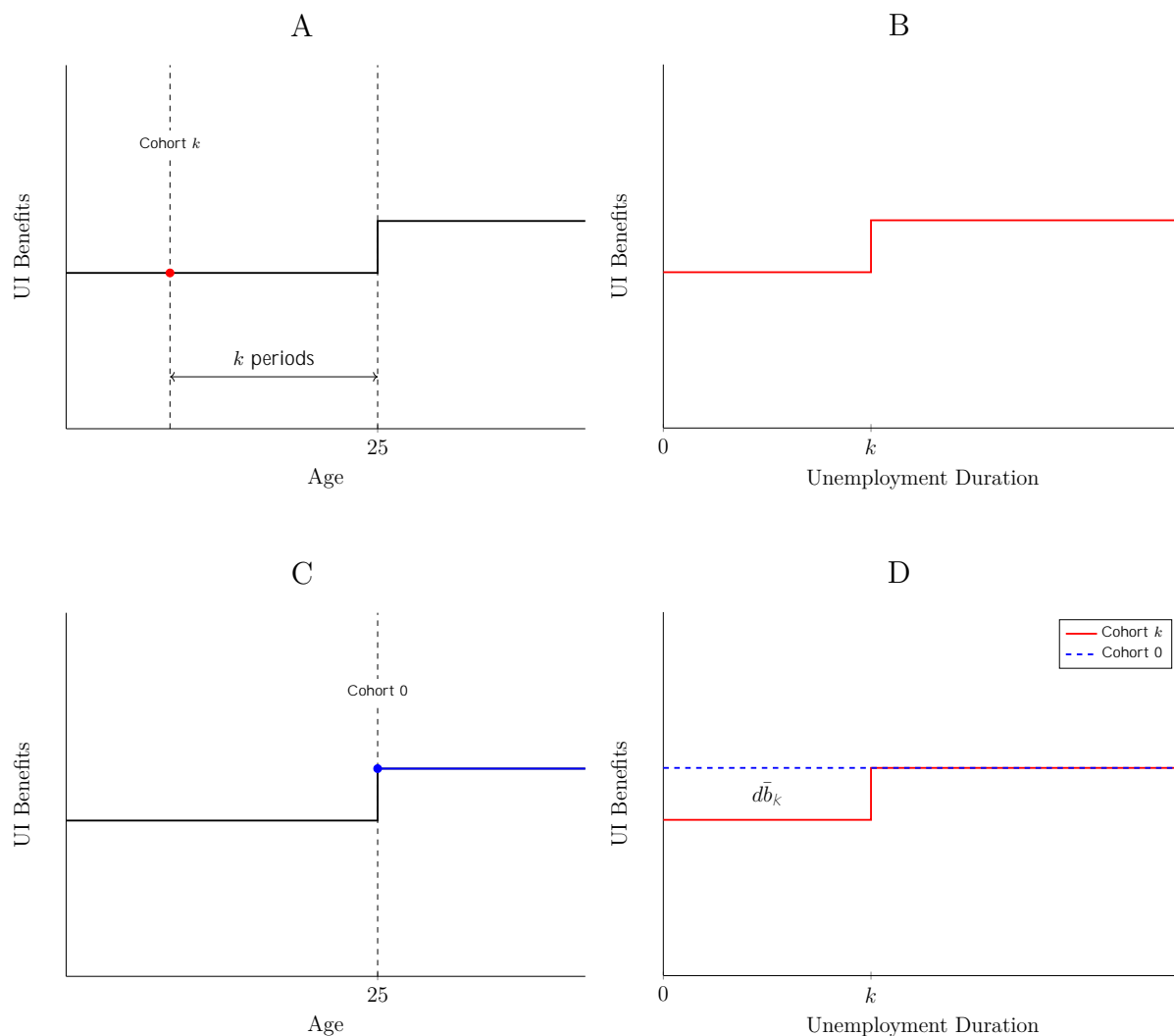
<sup>18</sup>Unfortunately, I have been unable to find a plausible explanation for the distortions in the distribution of claims. However, as this section shows, regardless of its underlying cause, this does not pose a threat to the identifying strategy implemented in this paper.

welfare cost of increasing UI benefits declines over the unemployment spell and providing more generous benefits becomes relatively less costly later in the spell. Given that existing evidence shows consumption smoothing gains of UI increase over the spell, my findings imply that introducing step-wise increases to the current benefit profile will improve welfare.

In order to explain why moral hazard costs of UI decline in the UK, in spite of the theoretical prediction to the contrary, I empirically examined the two main forces that influence the evolution of these costs over the spell. I found evidence of seemingly myopic search behaviour and non-stationarities in unemployment dynamics. These deviations from the assumptions of the theoretical job search models explain the discrepancy between theory and my empirical findings.

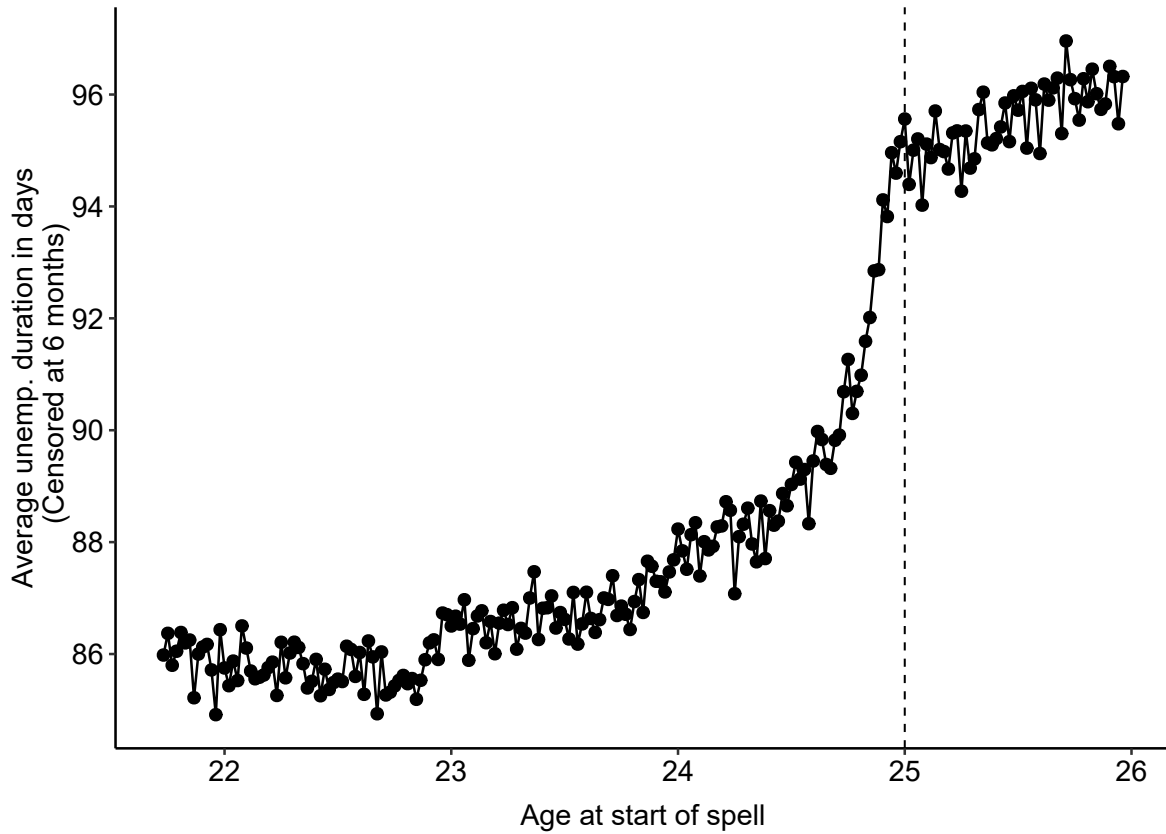
## 1.9 Figures and Tables

**Figure 1.1:** Variation in Time Profile of UI Benefits



**Notes:** The figure shows the variation in time profile of UI benefits. Panel A shows the benefit level around 25 years of age and marks cohort  $k$  as individuals who start a spell  $k$  periods before turning 25. Panel B shows the benefit profile for cohort  $k$ . Panel C shows the benefit profile (the blue segment) for cohort 0 who start their spell immediately after turning 25. Panel D shows how comparing the benefit profile of the two cohorts creates variation in benefits of the first  $k$  periods.

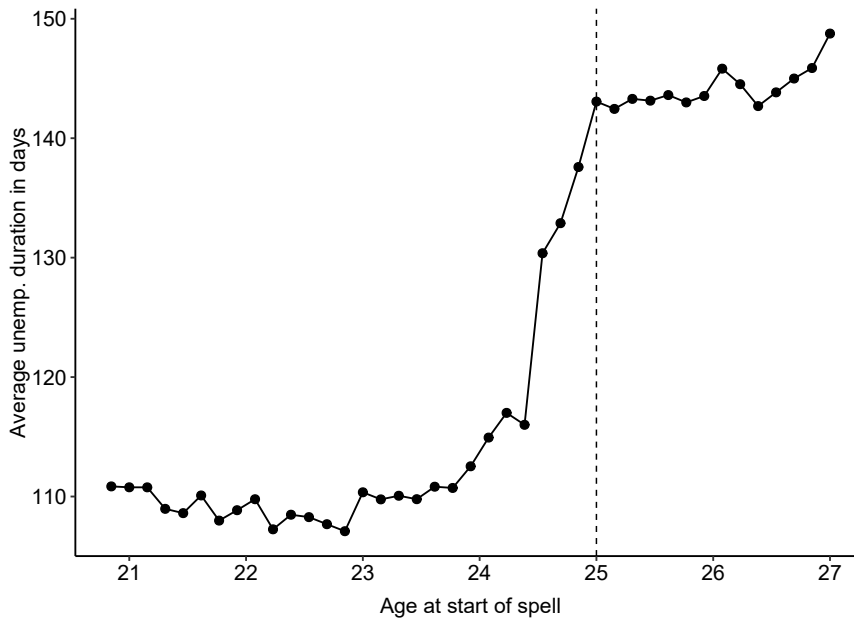
**Figure 1.2:** Average Censored Durations of Unemployment



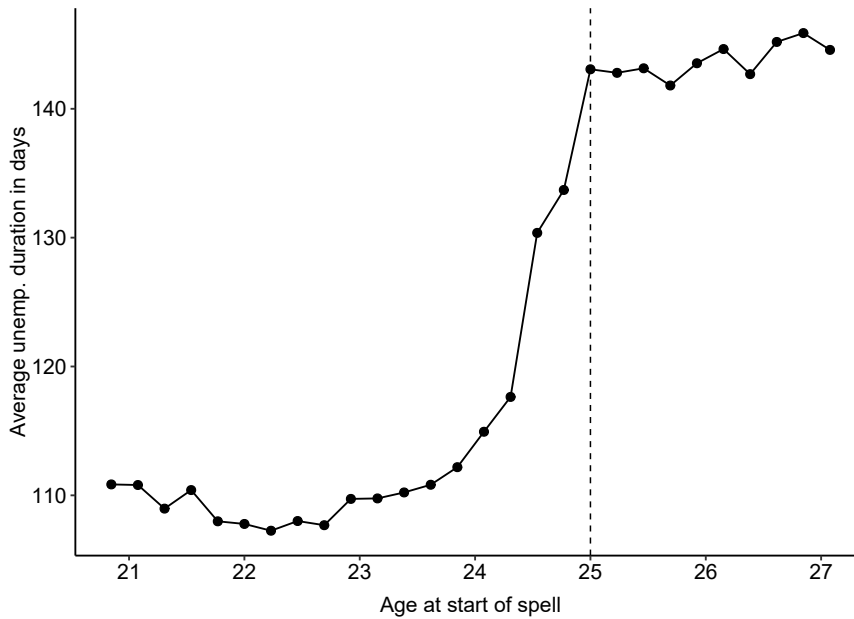
**Notes:** The figure shows the duration of unemployment censored at 6 months. The censoring is done to avoid the effect of the New Deal for Young People program (see section 1.3). The figure shows that duration of unemployment is stable immediately after 25 and for spells starting sufficiently earlier than 25. However, as claimants approach 25, the duration of unemployment steadily increases. Moreover, since UI benefit rises at 25, the curvature of this increase indicates that unemployment duration responds more strongly to changes in benefits paid earlier rather than later in the spell.

**Figure 1.3:** Average Duration of Unemployment

A. 8-week periods

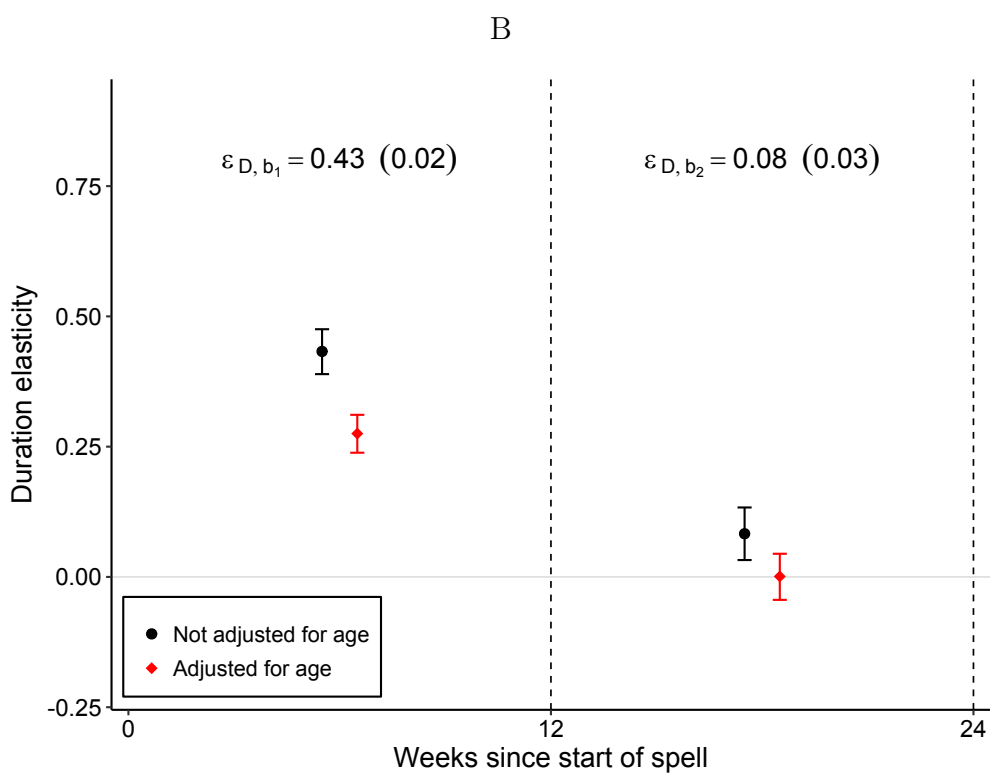
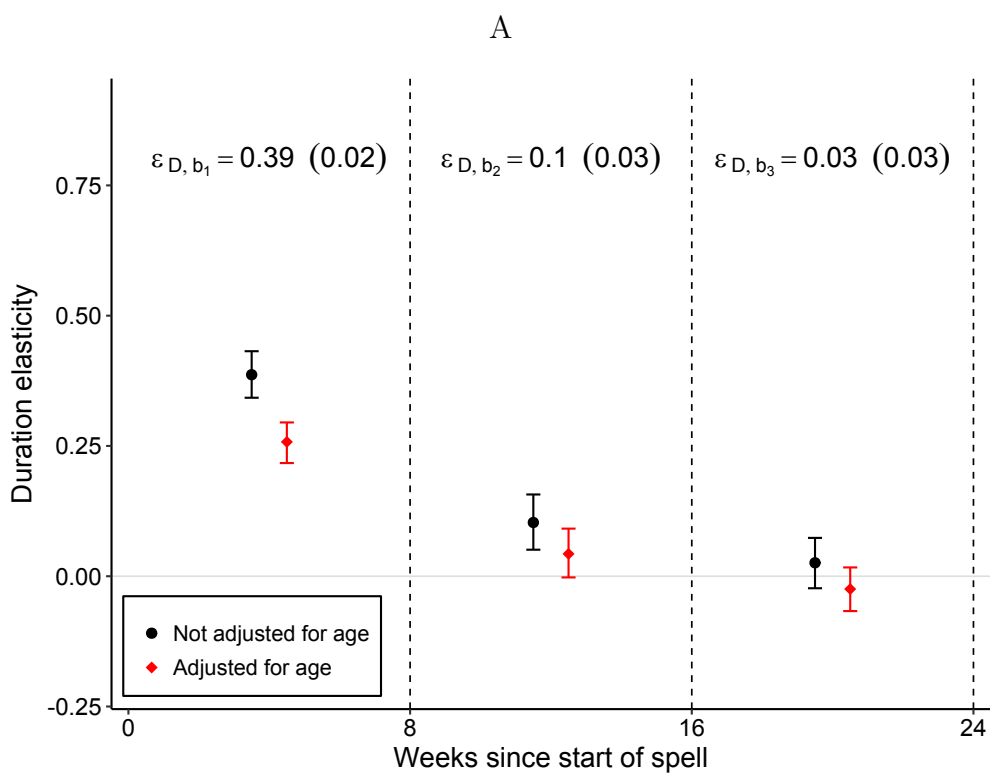


B. 12-week periods

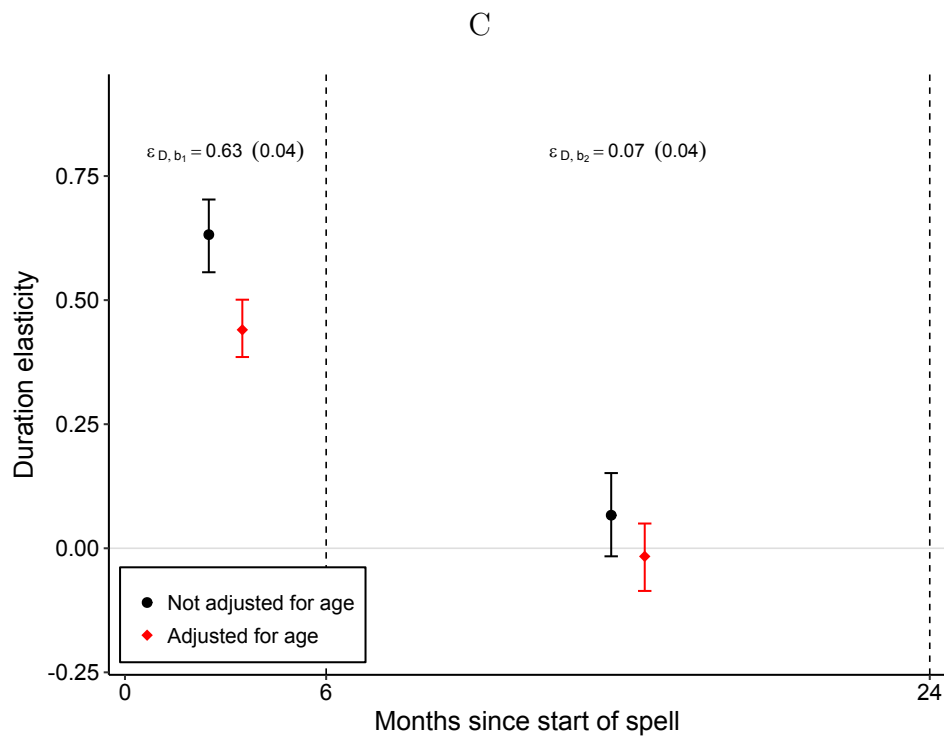


**Notes:** The figure shows the average duration of unemployment as a function of age for spells that start every 8 weeks and 12 week in panels A and B respectively. Unemployment duration is stable immediately after 25 and sufficiently earlier than 25. But it rises steadily as individuals approach 25. Moreover, since UI benefit rises at 25, the curvature of this increase indicates that unemployment duration responds more strongly to changes in benefits paid earlier rather than later in the spell. The sharp rise in durations between 25 and 25 is due to the New Deal for Young People program (see section 1.3) and is accounted for in the estimation.

**Figure 1.4:** Elasticity of Duration w.r.t Benefits Paid at Different Spell Lengths.

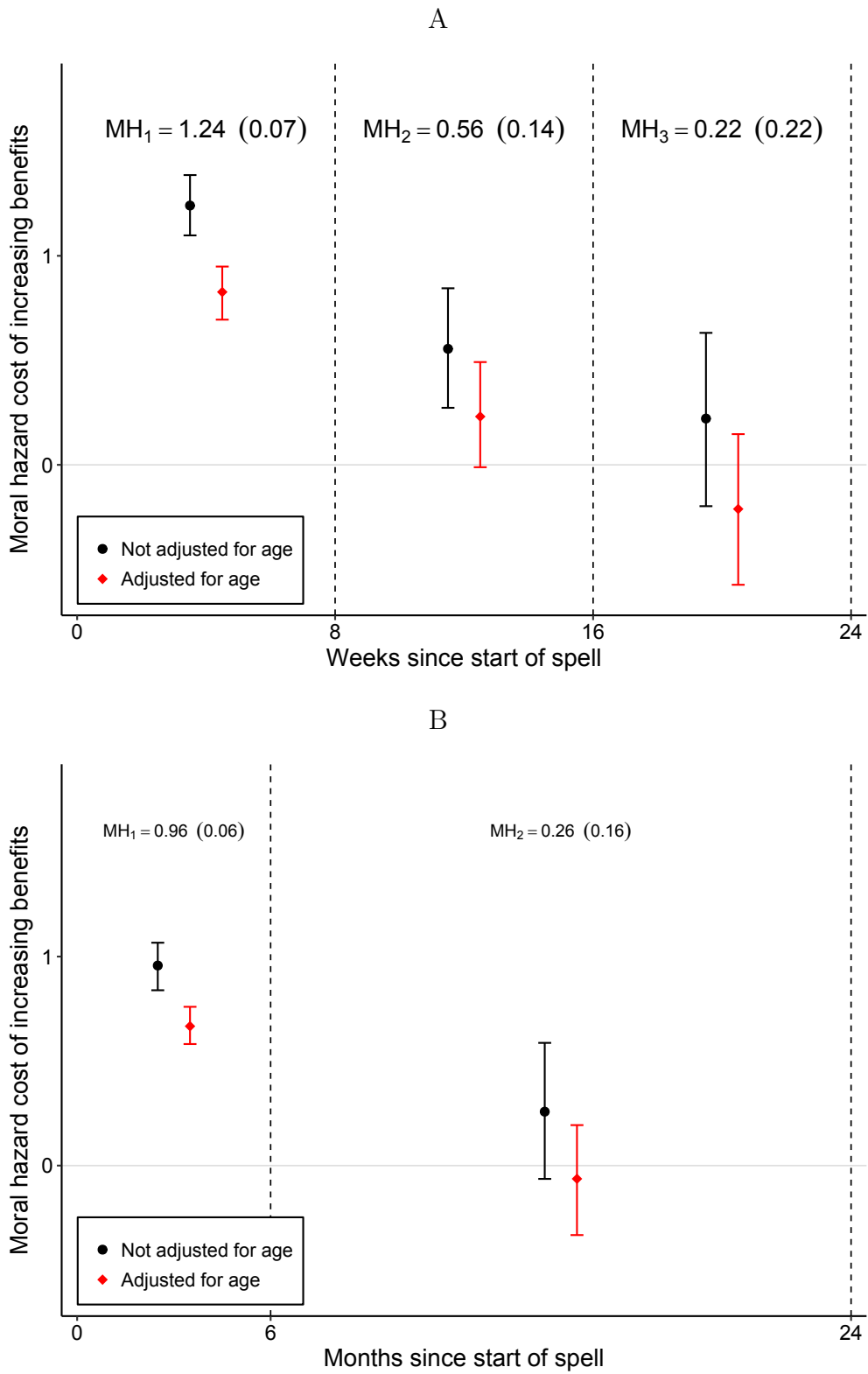


**Figure 1.4:** Elasticity of Duration w.r.t Benefits Paid at Different Spell Lengths (continued)



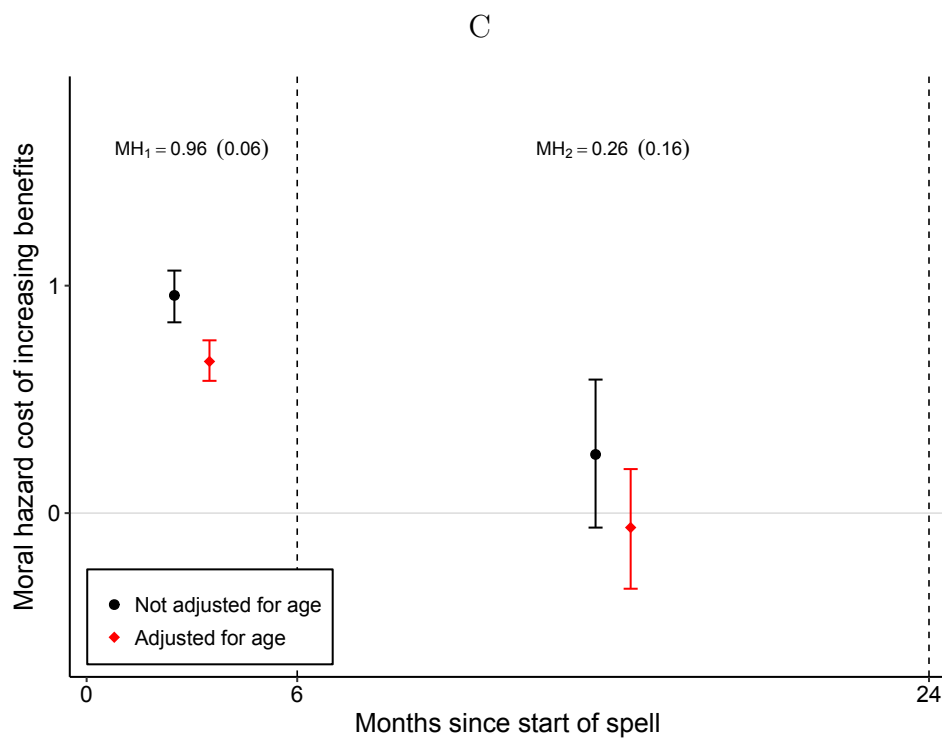
**Notes:** The figure shows estimates of elasticity of total unemployment duration w.r.t benefits paid at different parts of spell. Panel A and B report elasticities w.r.t 8-week and 12-week periods during the first 6 months, using spells from 2000-2015. The estimation is restricted to the first 6 months to avoid the effects of NDYP (see section 1.3). Panel C uses spells starting after 2011, when NDYP is no longer in effect, and reports elasticities w.r.t benefits paid in the first 6 months (26 weeks) and the next 18 months. Standard errors are computed by bootstrapping using 1000 replications.

**Figure 1.5:** Moral Hazard Cost of UI Benefits Paid at Different Spell Lengths



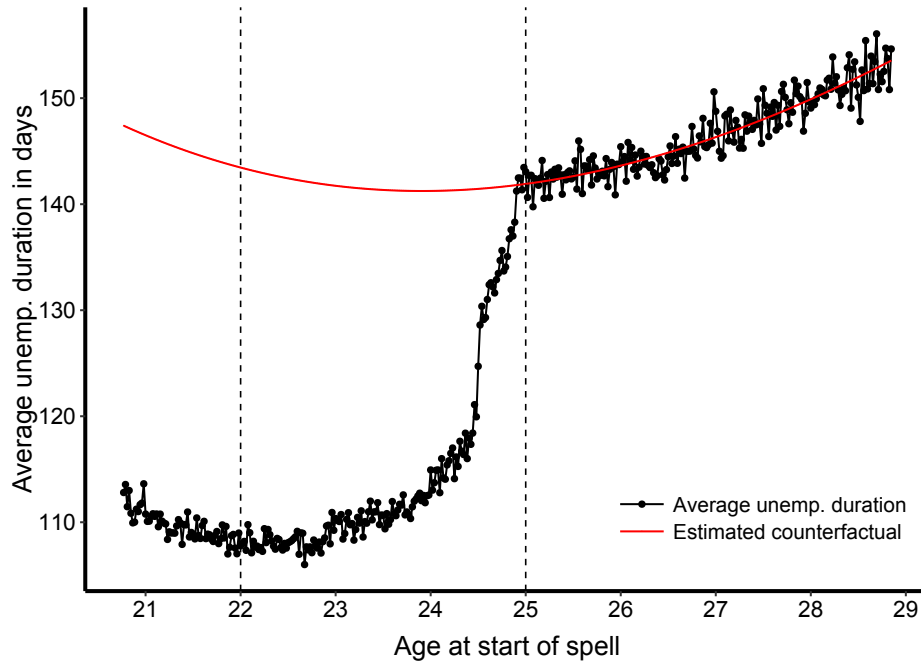


**Figure 1.5:** Moral Hazard Cost of UI Benefits Paid at Different Spell Lengths (continued)



**Notes:** The figure shows moral hazard cost of increasing benefits for different parts of spell. Panels A and B show the moral hazard cost for 8-week and 12-week long periods for the first 6 months, using spells from 2000-2015. The estimation is restricted to the first 6 months to avoid the effects of NDYP (see section 1.3). Panel C uses spells starting after 2011, when NDYP is no longer in effect, and reports moral hazard cost of increasing benefits for the first 6 months and for the next 18 months. Standard errors are computed by bootstrapping using 1000 replications.

**Figure 1.6:** Controlling for the Effect of Age on Unemployment Duration

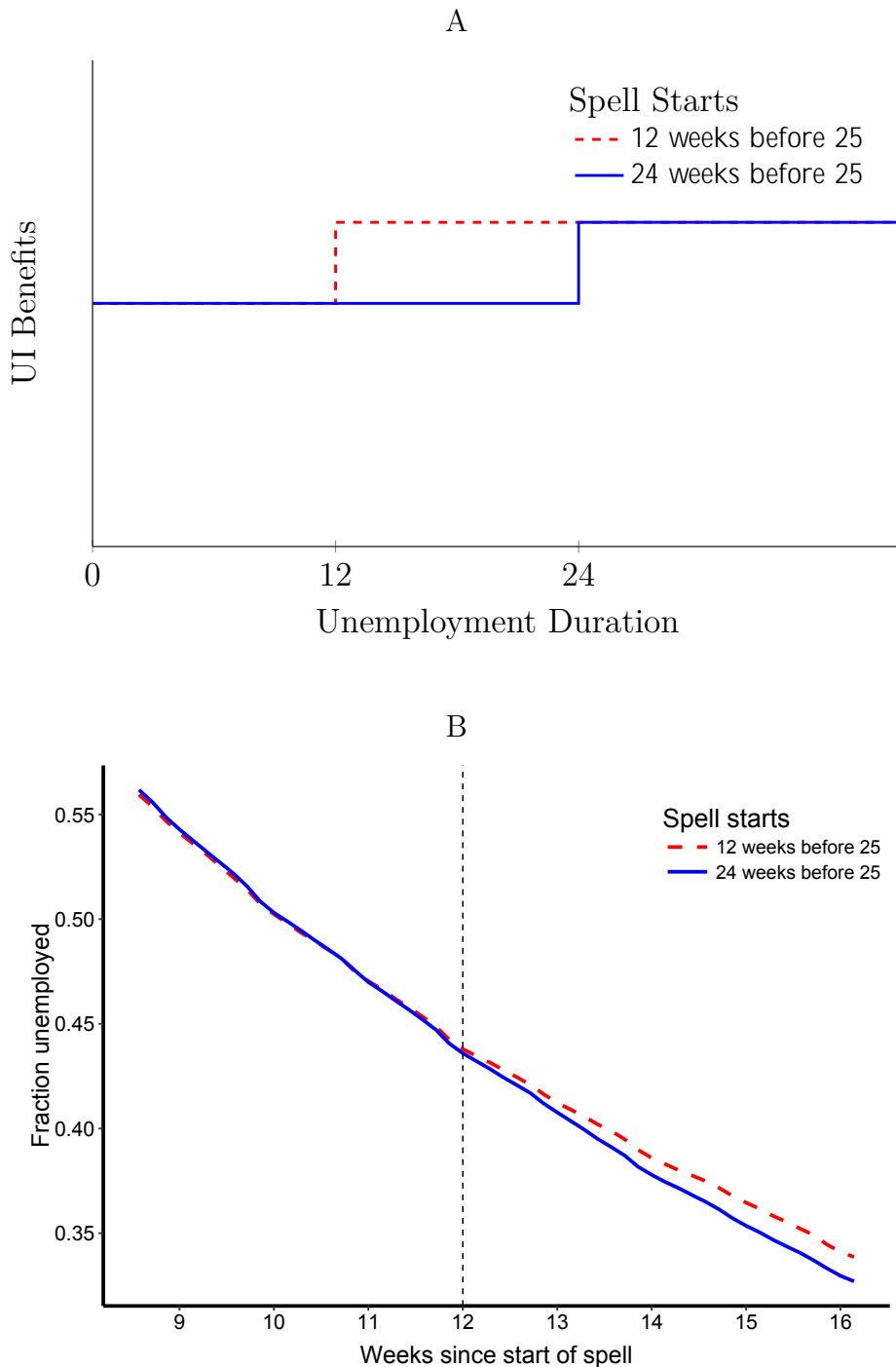


**Notes:** The solid red line shows the predicted relationship between age and unemployment duration if all individuals faced a flat benefit profile at the higher (25+) rate. To control for the possible direct effect of age on unemployment duration, I estimate the counterfactual relationship between age and duration in the absence of variation in benefits and adjust durations accordingly (see section 1.4). To estimate this counterfactual, I exclude spells starting up to 3 years before 25 and fit a flexible polynomial to the remaining observed durations:

$$D_i = \sum_{j=0}^q \beta_j \cdot (w_i)^j + \alpha \cdot I \{age_i < 22\} + u_i$$

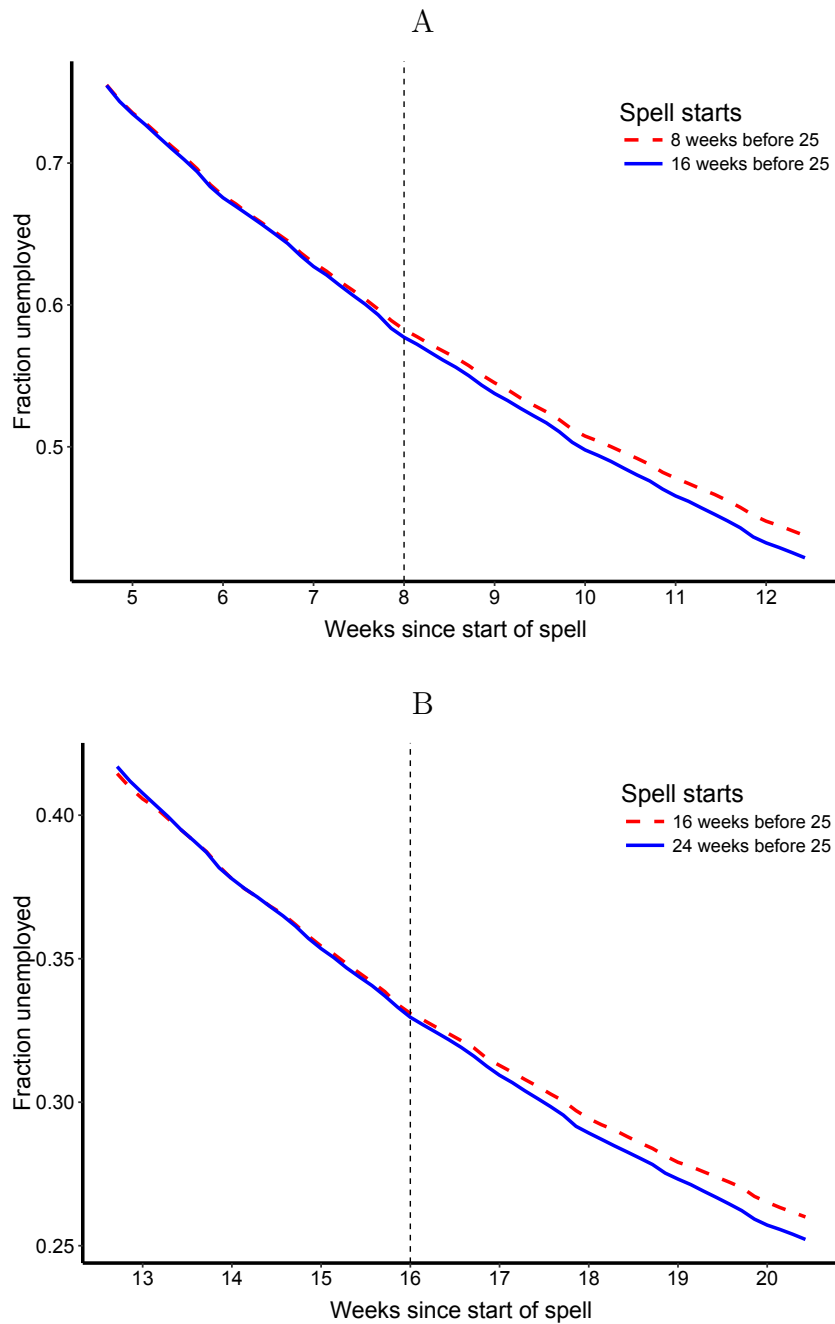
where  $\alpha$  captures the difference in durations between younger (under 22) and older (25+) claimants due to the NDYP program (see section 1.3) as well as the effect of facing different (constant) benefit levels.

**Figure 1.7:** Survival Curves for Cohorts Starting Spells 12 and 24 Weeks Before 25



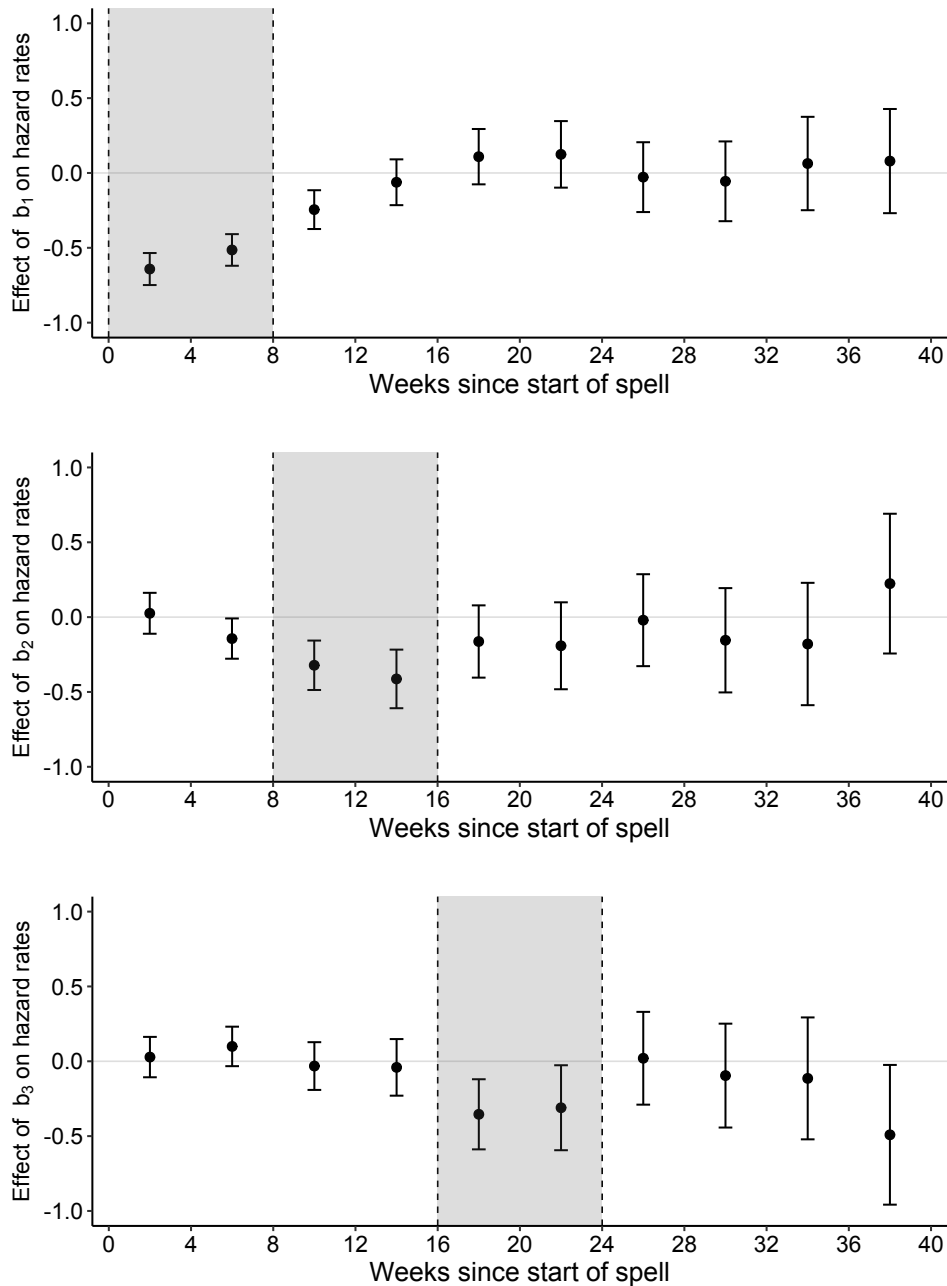
**Notes:** Panel A shows the UI benefit profile for two cohorts who start a spell 12-weeks and 24-weeks before turning 25. The survival curves in panel B show the fraction of job seekers still unemployed over time. With forward-looking job-seekers, we would expect the job finding rate of the 12-week cohort, who anticipate higher benefits in weeks 13 to 24, to be lower from the beginning of the spell. However, as panel B shows, the survival curves of the two cohorts are indistinguishable right up to week 12, when the 12-week cohort actually start receiving the higher benefits. It is only then that the job finding rates fall and the two survival curves diverge. This indicates that job-seekers do not react in advance to changes in future benefits.

**Figure 1.8:** Survival Curves for Cohorts Starting Spells 8, 16 and 24 Weeks Before Turning 25



**Notes:** The figure shows fraction of job seekers still unemployed as a function of unemployment duration. Panel A compares the survival rates of the cohorts who start their spell 8 and 16 weeks before turning 25 while panel B does the same for cohorts starting 16 and 24 weeks before 25. In both cases, the job finding rates are indistinguishable while the two cohorts receive the same amount of benefits. However, once the older cohort (solid blue) reach 25 and actually receive the higher benefit, their job finding rate decreases. This shows that search behaviour is not responsive to changes in future benefits and is mostly driven by current benefits.

**Figure 1.9:** Effect of change in current and future benefits on hazards out of unemployment.

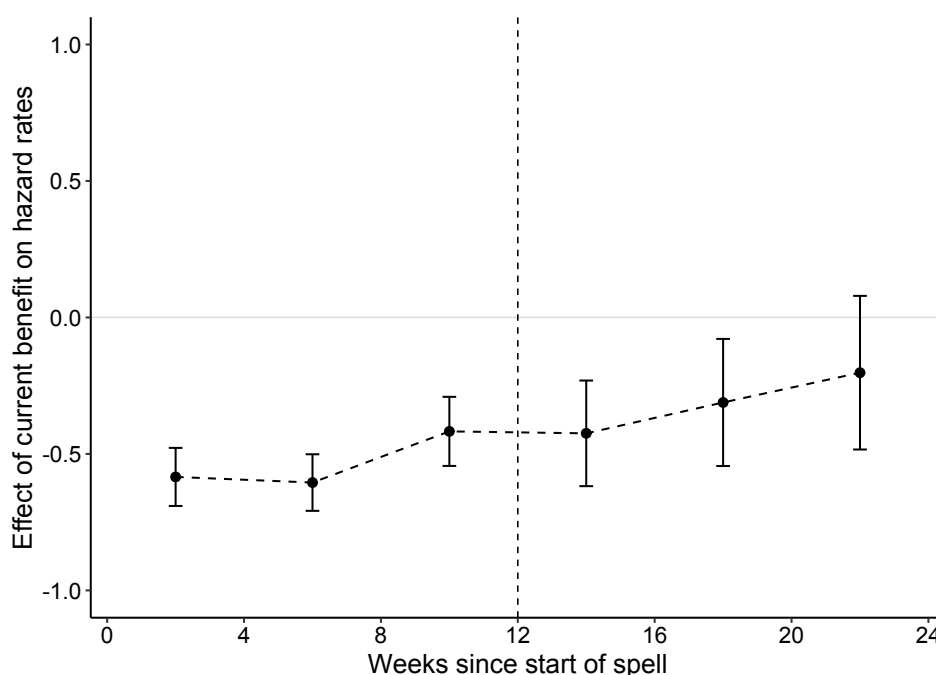


**Notes:** The figure shows the effect of a change in benefits paid in 8-week-long periods on exit rates from unemployment. The shaded interval in each panel marks the interval in which the corresponding benefits are changing. It clearly shows that exit rates respond to changes in contemporaneous benefits, but not to changes in future benefits, as nearly all estimates prior to the shaded intervals are insignificant. The effect of benefits in part  $k$  of spell,  $b_{k,t}$ , on job finding rates at time  $t$ , is estimated using proportional hazards models of the following form for individuals in cohorts  $k$  and  $k - 1$ :

$$\log h_{i,t} = \alpha_t + \beta_{k,t} \log b_{k,i}$$

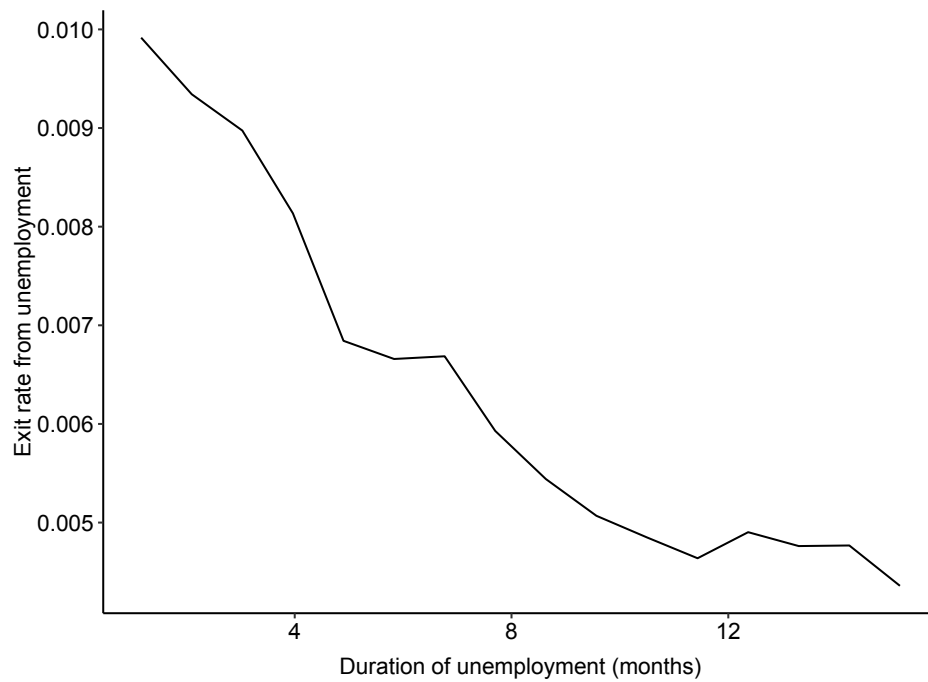
where  $h_{i,t}$  is the exit hazard for individual  $i$  at time  $t$  and  $b_k$  is the level of benefit for individual  $i$  in part  $k$  of spell.

**Figure 1.10:** Effect of changes in current benefits on exit rates from unemployment.



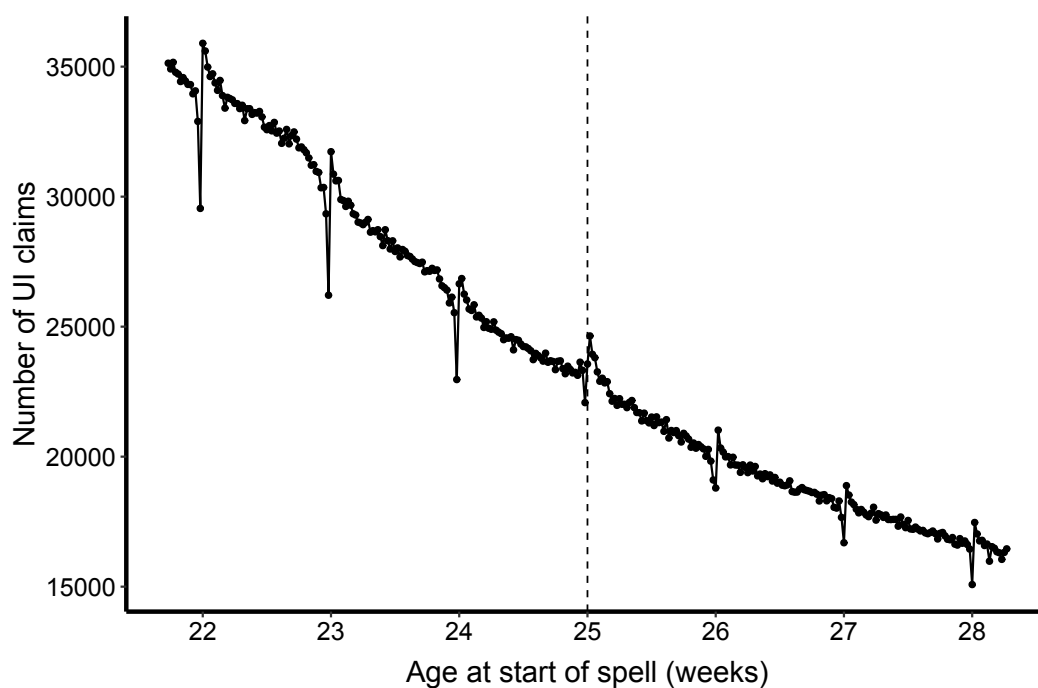
**Notes:** The figure shows the effect on the hazard of leaving unemployment over 4-week intervals of an increase in contemporaneous UI benefits paid over 12-week periods. That is, the three estimates on the left correspond to the effect of increasing benefits of the first 12 weeks and the three estimates on the right show the effect of increasing the benefits of weeks 12 to 24. Although benefits have a significant effect on probability of leaving unemployment early in the spell, the magnitude of this effect gradually declines towards zero. This indicates that unemployment dynamics are non-stationary and job finding rates become less responsive to benefits at longer durations.

**Figure 1.11:** Exit Rate From Unemployment



**Notes:** The figure shows exit rate from unemployment as a function of duration of unemployment for spells that start just after turning 25. In a stationary environment, we would expect the exit rates to be constant and independent of unemployment duration, as job seekers face a flat benefit profile after 25. However, the exit rate decreases steadily as unemployment durations increases, indicating that non-stationary forces (heterogeneity or true duration dependence) affect the probability of finding a job.

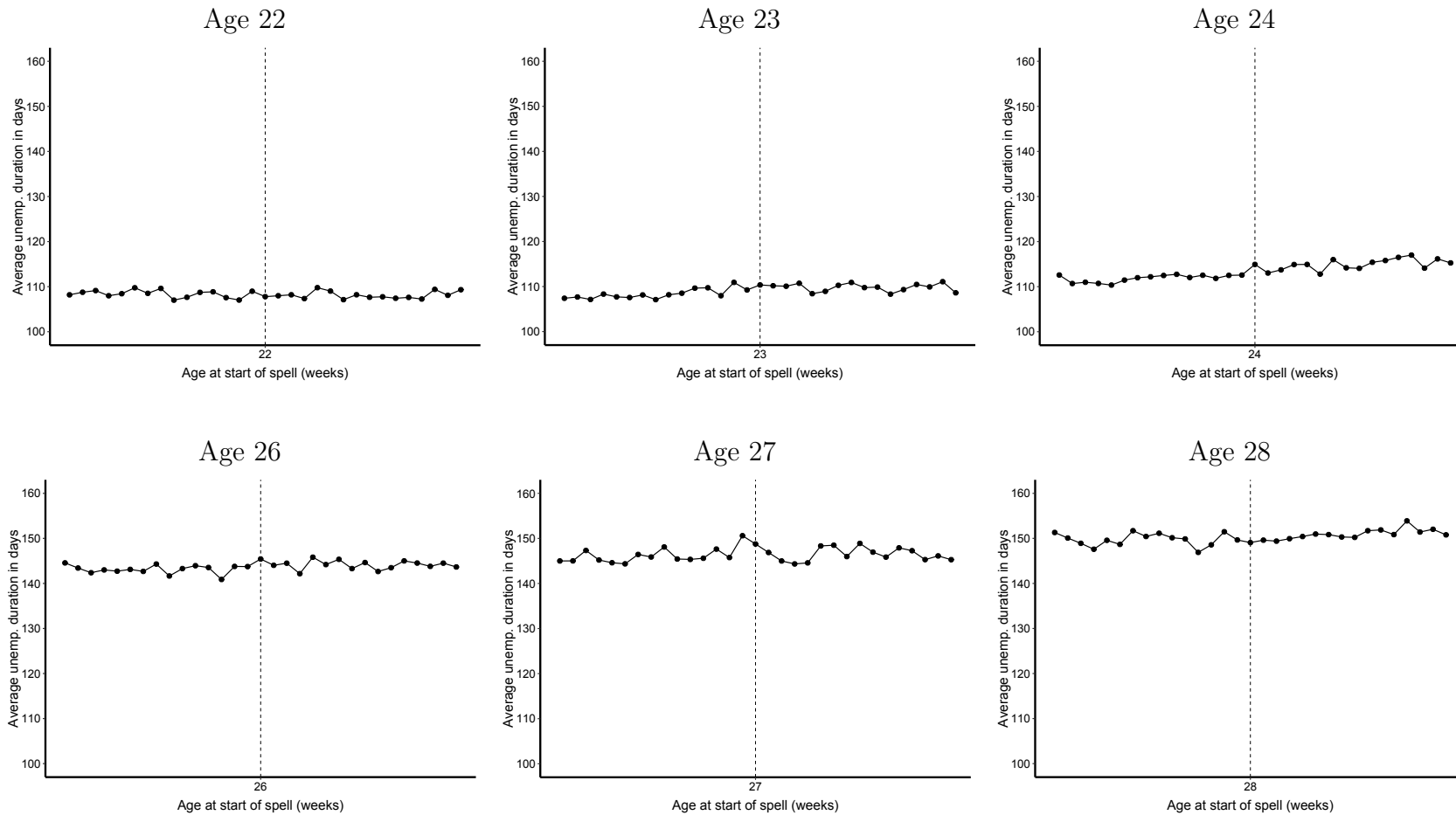
**Figure 1.12:** Number of UI Claims in Weekly Bins of Age



**Notes:** The figure shows the number of UI claims in weekly bins of age. The vertical line marks the week of 25-th birthday. Although there is a distortion in the distribution of claims around 25, the same distortions happen around all other birthdays. Given that there is no change in financial incentives (e.g. housing benefit, minimum wage, etc.), the distortions around 25 cannot be driven by financial incentives. section 1.7 shows that regardless of its nature, this distortion does not undermine identification of duration responses to benefits.



**Figure 1.13:** Duration of Unemployment Around Birthdays Other Than 25



**Notes:** The figure shows the average duration of unemployment in weekly bins around birthdays other than 25. Vertical lines mark the week of the corresponding birthday. The average duration evolves smoothly around all birthdays and shows no sign of a "birthday effect". This supports the identifying assumption that duration of unemployment is not affected by distortions in the distribution of claims around birthdays.

**Table 1.1:** Duration Elasticities and Moral Hazard Costs

	(A)		(B)		(C)	
	8-week periods		12-week periods		6 vs. 18 months	
	(1)	(2)	(1)	(2)	(1)	(2)
$\epsilon_{D,b_1}$	0.387*** (0.023)	0.258*** (0.020)	0.433*** (0.022)	0.275*** (0.019)	0.632*** (0.037)	0.440*** (0.030)
$\epsilon_{D,b_2}$	0.103*** (0.027)	0.043* (0.024)	0.083*** (0.026)	0.001 (0.022)	0.067 (0.041)	-0.016 (0.034)
$\epsilon_{D,b_3}$	0.026 (0.025)	-0.025 (0.021)				
$MH_1$	1.241*** (0.075)	0.827*** (0.064)	1.040*** (0.053)	0.661*** (0.046)	0.957*** (0.057)	0.667*** (0.045)
$MH_2$	0.555*** (0.145)	0.231* (0.128)	0.419*** (0.130)	0.005 (0.113)	0.258 (0.159)	-0.063 (0.133)
$MH_3$	0.221 (0.216)	-0.211 (0.180)				
Adjusted for age	No	Yes	No	Yes	No	Yes

**Notes:** The first 3 rows of the table report estimates of elasticities of unemployment duration from equation (1.4). Elasticities with respect to benefits paid over 8-week and 12-week periods during the first 6 months of the spell are reported in panels (A) and (B). Panel (C) divides the first two years of the spell into a 6 month and an 18 month periods. Rows 4-6 report the corresponding moral hazard costs. Moral hazard costs and duration elasticities decrease over the spell across all specifications. Column (2) of each panel also controls for the effect of age on duration of unemployment. Standard errors are computed by bootstrapping using 1000 replications.

**Table 1.2:** Duration Elasticities With Respect to Current and Future Benefits

	8-week periods	12-week periods
$\epsilon_{D_1, b_1}$	0.194*** (0.013)	0.262*** (0.016)
$\epsilon_{D_1, b_2}$	0.005 (0.015)	-0.014 (0.017)
$\epsilon_{D_1, b_3}$	-0.026 (0.015)	

**Notes:** The table reports estimates of elasticity of  $D_1$ , duration of unemployment in the first period, with respect to benefits paid over current and future periods, for the first 6 months of spell. Although the contemporaneous elasticities on the first row are significant, those with respect to future benefits on rows 2 and 3 are all very small and statistically insignificant. These estimates show that, contrary to what we would expect from forward-looking job search, the time spent unemployed in the first period is not responsive to changes in future benefits. Standard errors are computed by bootstrapping using 1000 replications.

**Table 1.3:** Robustness to the Method of Estimating the Counterfactual

	(A)		(B)		(C)	
	8-week periods		12-week periods		6 vs. 18 months	
	(1)	(2)	(1)	(2)	(1)	(2)
$\epsilon_{D,b_1}$	0.264*** (0.018)	0.258*** (0.020)	0.291*** (0.017)	0.275*** (0.019)	0.432*** (0.031)	0.440*** (0.030)
$\epsilon_{D,b_2}$	0.058*** (0.022)	0.043* (0.024)	0.025 (0.019)	0.001 (0.022)	-0.226*** (0.035)	-0.016 (0.034)
$\epsilon_{D,b_3}$	-0.006 (0.021)	-0.025 (0.021)				
$MH_1$	0.848*** (0.058)	0.827*** (0.064)	0.699*** (0.042)	0.661*** (0.046)	0.654*** (0.048)	0.667*** (0.045)
$MH_2$	0.314*** (0.118)	0.231* (0.128)	0.128 (0.097)	0.005 (0.113)	-0.875*** (0.135)	-0.063 (0.133)
$MH_3$	-0.055 (0.177)	-0.211 (0.180)				
Method	Linear	Polynom.	Linear	Polynom.	Linear	Polynom.

**Notes:** The table shows that the falling pattern in duration elasticities and moral hazard costs is robust to how the counterfactual is estimated. It reports estimation results similar to Table 1.1. To check the robustness of the estimates to the method of controlling for the effect of age, column (1) of each panel controls for age by estimating a linear counterfactual, based only on spells that start after 25. Column (2) repeats the results from Table 1.1 where the counterfactual is estimated using a flexible third degree polynomial based on spells that start before 22 and after 25. As the table shows the results in the two columns of each panel are very close. More importantly, the general result that duration elasticities and moral hazard costs decline over the spell is preserved.

**Table 1.4:** Placebo Elasticities Estimated Around Birthdays Other Than 25

	Age 26		Age 27		Age 28		Age 29	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$\epsilon_{D,b_1}$	0.017 (0.025)	0.004 (0.025)	-0.018 (0.024)	-0.036 (0.024)	-0.029 (0.027)	-0.051** (0.026)	-0.017 (0.027)	-0.018 (0.027)
$\epsilon_{D,b_2}$	-0.024 (0.031)	0.007 (0.030)	-0.023 (0.029)	-0.055* (0.029)	0.015 (0.033)	-0.020 (0.030)	0.023 (0.032)	0.043 (0.032)
$\epsilon_{D,b_3}$	0.018 (0.030)		-0.050* (0.029)		-0.078** (0.031)		0.020 (0.032)	

**Notes:** The table reports placebo estimates of duration elasticities around birthdays other than 25. The placebo elasticities are estimated in the same way as those around 25, using equation (1.4) and controlling for age using a third degree polynomial. Most estimates are insignificant. The few estimates that are statistically significant are very small and all have the wrong sign.

## Chapter 2

# Effects of Fiscal Stimulus: Evidence

## From the UK

### **2.1 Introduction**

During the Great Recession many governments around the world introduced fiscal stimulus packages to encourage economic activity in various markets, most notably the housing and the auto markets. These were usually in the form of a temporary subsidy or tax cut. Despite the consensus on the short-term effectiveness of such stimulus schemes, there is little agreement on whether they can have a longer lasting impact on the economy.

The short-term boost in market activity, during the period that the stimulus is in place, constitutes two different types of response. First, an extensive margin response by consumers who would not have entered the market otherwise, but do so due to the incentives offered by the scheme. This is what I will refer to as the long-term or the permanent impact of the stimulus, as this effect will not be reversed afterwards. Second, a timing response by consumers who merely shift their purchases forward in time to take advantage of temporary incentives offered by the scheme. Since these purchases would have otherwise happened in the near future and shortly after the end of the program, the re-timing response creates a drop in the volume of transactions once the stimulus scheme is over. This reversal drives a wedge between the size of the short-term and the long-term effects of such temporary incentive

schemes.

In this paper, I will first present estimates of the short-term and long-term effects of a tax cut in the UK housing market and show that a substantial part of the response to this tax cut has been through the extensive margin and it has therefore had a sizeable permanent effect. I will then argue that, in the presence of frictions in the market, the share of extensive margin response out of total short-term response is directly related to the duration of the incentive scheme. If the stimulus runs for a short period of time, most of the response will be driven by re-timing and will be reversed shortly after the program ends. This implies that a short-lived temporary incentive will not succeed in creating a considerable permanent effect, while a longer stimulus is likely to have a relatively larger permanent impact.

This paper is related to the body of work that provides evidence on the extent to which short-term timing response to taxes may exceed long-term responses (Auerbach and Poterba 1988; Burman and Randolph 1994; Goolsbee 2000) as well as the more recent literature on estimating the effects of fiscal stimulus using micro data, including Johnson, Parker, and Souleles (2006), Agarwal, Chunlin, and Souleles (2007) and Parker et al. (2013) who examine the effect of tax rebates on consumption of non-durables and find strong short-term effects but no evidence of reversal afterwards.

More closely related to this paper are Mian and Sufi (2012) and Best and Kleven (2017). Mian and Sufi (2012) find that the short-term effect of the 2009 CARS<sup>1</sup> program in the US was almost entirely driven by re-timing responses and was almost completely reversed due to lower market activity in subsequent months. They conclude that "although the initial effect of CARS on auto purchases is large, there is strong evidence of swift program reversal", and that almost all of purchases under the CARS program would have otherwise occurred within 10 months after it ended. On the other hand, Best and Kleven (2017) study the UK Stamp Duty holiday of 2009, a temporary transactions tax cut in the UK housing market, and find that nearly 70% of the short-term response was along the extensive margin and only 30% of it was reversed

---

<sup>1</sup>Cars Allowance Rebate System (CARS) program, commonly referred to as Cash for Clunkers.

during the year following the tax holiday.

In this paper I will try to take a first step towards reconciling these two seemingly contradictory results by arguing that, in markets with frictions, the share of extensive margin out of total short-term response, is linked to the duration of the program<sup>2</sup>. I begin by presenting estimates of the short-term and long-term effects of a temporary tax cut in the UK housing market showing that this temporary incentive has had a sizeable effect on market activity not just in the short term (i.e. during the stimulus) but also in the long-term. The second part of the paper will illustrate the link between the duration of a stimulus and its permanent impact using a simple search model of the housing market.

Section 2.2 will explain the context of the UK Stamp Duty reform of 2015 and estimate its short-term and long-term effects on market activity. Section 2.3 will present the model to illustrate the effect of program duration on its long-term impact. Section 2.5 Concludes.

## **2.2 The Effects of UK Stamp Duty Reform**

### **2.2.1 Context and Data**

Property transactions in the UK are liable to a transaction tax known as the Stamp Duty Land Tax (SDLT). Until April 2015, SDLT was applicable to transactions in all of the UK (England, Scotland, Wales and Northern Ireland). However, Scotland Act 2012 provided the Scottish Parliament with the power to introduce devolved taxes. As a result The Scottish Government announced in October 2014 that starting from April 2015 a new tax called Land and Buildings Transaction Tax (LBTT) will replace SDLT in Scotland.

In the meantime, on 4th December 2014, four months before LBTT was due to come into effect, the UK government announced an unanticipated reform to SDLT with immediate effect. This reform "undercut" Scotland's LBTT rates for a wide range of house prices. Notice that this unanticipated reform would apply to all of the UK, including Scotland where SDLT would remain

---

<sup>2</sup>I will assume throughout that the start of the stimulus program is unanticipated but its end date is anticipated, as was the case in UK Stamp Duty holiday, CARS in the US and the reforms that I study in this paper.



applicable for four months until April 2015. This meant that the Scottish housing market would experience two tax reforms in a short space of time: First, the unanticipated reform of the UK-wide SDLT in December 2014, followed by the pre-announced implementation of Scottish LBTT in April 2015.

Table 2.1 shows tax rates under old SDLT, new SDLT and LBTT. Figure 2.1 shows total tax liability under each tax regime. The three schedules are numbered in the chronological order of their applicability in Scotland. The combination of these reforms creates interesting variation in tax liability in Scotland over time, which I will exploit for the empirical analysis that follows.

I combine data from two sources. For the pre-LBTT period (i.e. before April 2015), I use administrative data on tax returns provided by HMRC covering all property transactions in the UK. This dataset contains rich information on tax return for each transaction, but has very little information otherwise. For the post-LBTT period, I use total number of monthly transactions in £5000 bins provided by Revenue Scotland.

## 2.2.2 Empirical Strategy

I employ a difference-in-differences strategy to estimate the short-term effect as well as the extent of reversal after the end of tax cut. I use the transactions in [£335K, £480K] as the treated group. Figure 2.1 shows that the transactions in this group are taxed less under the new SDLT. This means that these properties received a tax cut in December 2014 when SDLT was unexpectedly reformed. However, the same price range would be liable to a higher level of tax under Scotland's LBTT which was due in April. Therefore, in the [£335K, £480K] price bracket, the combination of these two reforms proxy a temporary and unanticipated tax cut. It should be noted that the upper bound of this bracket was chosen such that it would be sufficiently away from the £500K notch in the old SDLT schedule. Best and Kleven (2017) document that these notches induce considerable bunching. The SDLT reform removes the notch at £500K. This would reduce the number of transactions just below £500K as the extra mass around the notch gradually disappears. The treated price range is chosen such that it does not include the bunching region around £500K.

It is worth noting that not all properties in the treated price range receive

the same treatment. Properties in the lower parts of the treated price bracket receive a relatively larger cut followed by a smaller increase in tax rates than the ones with higher prices. I use this rather wide treatment range to improve statistical power. However, this affects the interpretation of the estimates. One could think of the estimated response as the response to the average changes in tax rates. Furthermore, the main purpose of the empirical exercise here is to show that the long-term response to such changes in tax rates is not zero rather than providing precise estimates of the magnitude of the response.

As the counterfactual, one would ideally use a group of transactions that are unaffected by these tax reforms and the transactions below £125K indeed provide this opportunity. But the market for properties under £125K is unlikely to be sufficiently similar to that of much more expensive properties in the treated range<sup>3</sup>.

The closest alternative is to use the transactions towards the top end of the first tax bracket. Although the reforms affect tax liability in this bracket but the magnitude of the changes are relatively small, especially for properties close to £250K. I use the [£200K, £230K] price range as the control group. Once again, this is chosen to exclude the bunching region close to the £250K Notch. The results are not sensitive to the exact width of this range.

To estimate the short-term effect and the reversal, I implement a difference-in-differences strategy using the following regression:

$$n_{it} = \alpha_0 Pre_t + \alpha_1 Cut_t + \alpha_2 Rev_t + \alpha_3 Post_t + \alpha_4 Treat_i + \beta_S Cut_t \times Treat_i + \beta_r Rev_t \times Treat_i + \beta_p Post_t \times Treat_i + u_{it} \quad (2.1)$$

where  $n_{it}$  is the log number of transactions for group  $i$  (treated and control) at time  $t$ , while  $Pre_t$ ,  $Cut_t$ ,  $Rev_t$  and  $Post_t$  are dummies for pre-reform, tax cut, reversal and post-reversal periods respectively.  $\beta_S$  and  $\beta_r$  will capture the short-term effect and the reversal respectively, while  $\beta_p$  will indicate whether the two series completely converge after the reversal period, implying that the reversal period is over. The duration of the reversal period is visually chosen

---

<sup>3</sup>Indeed, a closer look at the data shows that the two groups had been following different trends before the reform.

to be two months. However, the results are not sensitive to this choice.

### 2.2.3 Results

Figure 2.2 shows the normalised number of transactions in treated and control price brackets between April 2010 and Feb 2016. Each series is rescaled by its own mean. The two vertical dashed lines show the time period of interest. The first line marks December 2014 when the tax was cut because of the new SDLT rates introduced by the UK government. The second line indicates April 2015, when the higher Scottish LBTT rates were anticipated to come into effect.

First, Figure 2.2 shows that during the four preceding years the control series closely follows the treated series. The notable exception to this appears to happen in July and August 2014. The volatility in these two months might have been due to the uncertainty in the run up to the Scottish independence referendum, which took place in mid-September. Notice in particular that the control and the treated series remain parallel in the post-reform period. This alleviates the concern that the control group is receiving a small tax cut and the potential bias that this can introduce.

To see the response more clearly, Figure 2.3 plots the two series after January 2014. Bunching around the time notch and the missing transactions after the notch are clearly visible in this graph. Also, to make the comparison with the prediction of the model of the next section easier, Figure 2.4 plots the difference in activity between the treated price bracket and the counter-factual.

Figure 2.4 shows that in the months before the first reform, the difference between the two series is relatively stable around zero, except for the volatility before the Scottish independence referendum. After the unanticipated tax cut, there is a small increase in activity in the first two months. As we get closer to the end of the tax cut the volume of transactions increases rapidly and peaks in March, the last month before the end of the tax cut. However, the fact that the difference becomes negative and large for the months immediately after March, suggests that a considerable number of transactions were pulled forward in response to the tax cut. This subsequent drop in activity reverses the short-term effect of the tax cut.

Table 2.2 reports the results for Equation 2.1. The short-term effect cap-

tured by  $\beta_S$  shows that during the tax cut monthly activity has been, on average, 47 to 50 percent higher compared to what it would otherwise be. This accumulates over four months from December to March. However, the large and negative  $\beta_r$  shows that subsequent monthly activity has been 50 to 54 percent lower than what it would have been had there not been a tax cut. This will, to some extent, reverse the short-term effect over the two-month reversal period. The bottom row of the table calculates the share of permanent effect out of the total short-term response, as  $1 - \frac{2|\beta_r|}{4\beta_S}$ . These estimates imply that up to 60 percent of the increase in the number of transactions during the tax cut can be attributed to extensive margin response, and will therefore be a permanent effect.

The next section will present a simple search model to highlight the link between the duration of a temporary stimulus and the magnitude of its permanent effect.

## 2.3 Model

This section presents a search model of the housing market. The aim of this simple model is to highlight the idea that the share of the permanent effect of a temporary stimulus in its total short-term effect is directly related to its duration.

The intuition behind the model is as follows. In deciding whether to enter the market during the tax cut, individuals consider how long it is left until the stimulus period is over. When entering the market at later stages (i.e. closer to the end of stimulus) individuals are less likely to find a match before the end. Therefore, the closer we get to the end the fewer people will find it worthwhile to start searching. This means that the extensive margin response (i.e. number of new entries) will fall towards the end of the stimulus. A short-lived stimulus program will not provide enough time for individuals to find a match and will be less successful in persuading people to enter the market. In contrast a longer stimulus program will provide more time for searching and will on average induce a larger number of people to enter the market. This will in turn increase the permanent effect of the program.

### 2.3.1 Setup

Every period a new cohort of individuals are born and decide whether to enter the market and search for a match or take no action. Let  $F(\cdot)$  denote the cumulative distribution of the present discounted value of inaction for each cohort and  $V_t$  the value of entering the market in period  $t$ . Therefore, the fraction  $F(V_t)$  of cohort  $t$  will enter the market as the value of doing so is larger than the value of taking no action for them. To simplify the analysis, I will assume that once individuals enter, they will not be able to dropout of the market and have to search until a match is found.

Individuals who have entered the market, choose their search effort  $s_t$  every period, which is normalised to equal the probability of finding a match. Cost of effort is given by the cost function  $c(s)$ , with  $c(s), c'(s) > 0$ . Let  $H$  denote the present discounted value of the stream of utilities from finding a match. The price of the commodity exclusive of tax,  $P$  is assumed to be constant. Let  $B(\tau) = H - P(1 + \tau)$  denote the value of a match given tax rate  $\tau$ .

The model abstracts from the supply side by assuming that pre-tax prices are fixed. This may not be a precise description of some markets, including the housing market where the supply is rather inelastic, especially in the short run. However, this assumption simplifies the model considerably and a similar one-sided model of the supply side can be shown to lead to similar results.

The number of people searching at time  $t$ , denoted by  $A_t$ , evolves according to  $A_t = A_{t-1}(1 - s_{t-1}) + F(V_t)$ . The number of matches formed in the market in period  $t$  is given by  $N_t = A_t s_t$ . Combining the two yields:

$$N_t = (A_{t-1}(1 - s_{t-1}) + F(V_t))s_t \quad (2.2)$$

This demonstrates that to characterize the dynamics of the number of transactions we need to understand the paths of  $V_t$ , the value of entering the market and  $s_t$ , agents' search effort. This will then shed light on how the duration of the tax cut affects the share of extensive margin response out of the total short-term response.

### 2.3.2 Extensive Margin Response

Consider an unanticipated and temporary tax cut which reduces the tax rate from  $\bar{\tau}$  to  $\underline{\tau}$ . The tax cut starts at  $t = 0$  and ends at  $t = T$  at which point the tax rate resumes to  $\bar{\tau}$ . The end of the tax cut is announced in advance and is known to individuals.

The value of searching at time  $t$  is given by:

$$V_t = \max_{s_t} \{s_t B(\tau_t) + \theta(1 - s_t)V_{t+1} - c(s_t)\} \quad (2.3)$$

where  $\theta$  is agents' discount factor.

When the tax rate is not expected to change individuals face the same continuation value every period and  $V_t = V_{t+1} = V(\bar{\tau})$ . This implies that in the steady state with  $\tau = \bar{\tau}$ , the same fraction,  $F(V(\bar{\tau}))$  of every new cohort enter the market and start to search. This is true before the start of the tax cut at  $t = 0$  and also for  $t = T$  when the tax cut is over. During the temporary tax cut

$$V_t = \max_{s_t} \{s_t B(\underline{\tau}) + \theta(1 - s_t)V_{t+1} - c(s_t)\}, \quad 0 \leq t < T. \quad (2.4)$$

Notice that  $V_t = V_{t+1}$  during the tax cut as agents expect the tax rate to change in the future. To characterize the dynamics of  $V_t$ , let  $h(x) = \max_s \{sB(\underline{\tau}) + \theta(1 - s)x - c(s)\}$  be the function that relates  $V_{t+1}$  to  $V_t$ . Using this function we can rewrite equation (2.4) as

$$V_t = h(V_{t+1}), \quad 0 \leq t < T. \quad (2.5)$$

Notice that  $h(\cdot)$  is a contraction<sup>4</sup>. Therefore, contraction mapping theorem implies that the iterative sequence  $\{V_t\}$  will monotonically converge to  $V(\underline{\tau})$ , (which is the fixed point of  $h(\cdot)$ ), as  $t \rightarrow -\infty$ , i.e. as we move backwards in time towards the beginning of the tax cut. For ease of notation, let  $n$  denote the number of periods left until  $T$ , so that  $t = T - n$ . We can then rewrite (2.5) as  $V_{T-n} = h(V_{T-(n-1)})$ . This implies that the sequence  $\{V_{T-n}\}$  monotonically

---

<sup>4</sup>That is,  $h(\cdot) = \theta(1 - s) \in [0, 1)$ .

converges to  $V(\underline{\tau})$  as  $n \rightarrow \infty$ , i.e., as we move away (backwards in time) from the end of the tax cut. Furthermore, since the sequence  $\{V_{T-n}\}$  starts at  $V(\bar{\tau})$  and monotonically converges to  $V(\underline{\tau})$ , and  $V(\underline{\tau}) > V(\bar{\tau})$ , it must be increasing in  $n$ . Remember that  $n$  is the number of periods until the end of tax cut. This means value of entering the market,  $V$ , is larger in earlier periods of the tax cuts that are further away from the end of tax cut. In other words, the value of entering the market, and therefore the fraction of individuals who decide to enter, is higher in periods that are further away from the end of the tax cut.

Figure 2.6 shows the evolution of  $V_t$ , the value of entering the market over time. It demonstrates why a short tax cut will create limited extensive margin response. During a short tax cut the value of entering the market is relatively low and fewer people will do so. But as the tax cut gets longer the value of entering the market and starting to search rises towards  $V(\bar{\tau})$  and more people will enter the market in earlier period. This means that the average number of new entries per period will rise with the duration of the tax cut.

### 2.3.3 Search Effort

The second component that affects the dynamics of the response to the tax cut is search effort. This section shows how agents' search effort responds to this temporary incentive.

Agents choose search effort  $s_t$  by solving the optimization problem in equation (2.3). The F.O.C. for this problem is:

$$c(s_t) = B(\tau) - \theta V_{t+1} \quad (2.6)$$

In a steady state when the tax rate is not expected to change,  $V_{t+1} = V_t$  and agents will choose the same  $s_t$  every period. Let  $\bar{s}$  and  $\underline{s}$  denote the steady state search effort when the tax is  $\bar{\tau}$  and  $\underline{\tau}$  respectively.

During a temporary tax cut, given that  $c(\cdot)$  is an increasing function, equation (2.6) implies that  $s_t$  will mirror the path of  $V_t$  described in subsection 2.3.2. That is,  $s_t$  will be highest in the last period before the end of tax cut and will monotonically converge to  $\underline{s}$  as  $t \rightarrow -\infty$ , i.e. as we move to earlier periods of the stimulus.

Figure 2.7 shows the evolution of search effort during a temporary tax cut. The figure shows that search effort monotonically increases during the cut period and reaches its highest level just before the end. However, once the tax cut is over it falls back to its high-tax steady state level,  $\bar{s}$ .

Intuitively, the search effort rises towards the end of the tax cut because individuals are trying to find a match while the incentive is still in place. This rise in search effort is the force that drives the re-timing response to the stimulus.

## 2.4 Total Response And The Share of Extensive Margin

Having characterized the dynamics of the extensive margin response and search effort, we can put these together in equation (2.2) and numerically simulate the number of transactions every period. Figure 2.8 shows the simulated number of transactions before, during and after the temporary tax cut. As this figure shows, the number of transactions increases sharply in the periods leading up to the end of the tax cut and falls below its high-tax steady state level immediately afterwards, before gradually converging to it. The former increase is driven by the increase in search effort, as individuals try to find a match before the tax increases at  $t = T$ . This means that a large number of individuals leave the market just before  $t = T$  which, combined with the sudden drop in search effort after  $t = T$ , results in a sharp drop in the number of matches just after  $t = T$ .

In the absence of the stimulus incentive, the steady state number of transactions would have been constant and equal to  $F(V(\bar{\tau}))$ , shown by the dashed line in Figure 2.8. Let  $N^C = F(V(\bar{\tau}))$  represent this counterfactual level. The short-term effect of the stimulus is given by the number of transactions in excess of  $N^C$  between  $t = 0$  and  $t = T$ . The permanent effect of the stimulus is equal to the number of individuals who enter the market that would not have done so in the absence of the tax cut, that is  $\sum_{t=0}^{t=T} (F(V_t) - N^C)$ . In other words, the area under  $F(V_t)$  and above  $N^C$  gives the permanent effect of the stimulus.



Figure 2.9 shows the number of transactions and the number of entries into the market in excess of the counterfactual level. Remember that the area under the former between  $t = 0$  and  $t = T$  gives the short term response while the area under the latter is equal to the long-term effect. This figure reveals how the duration of the tax cut affects the share of extensive margin out of the short-term response. In particular, it shows that the response to a short-lived stimulus will be dominated by the sharp rise in transactions driven by increased search effort while creating a limited extensive margin response. However, as the duration of the stimulus increases,  $V_t$  rises (once again holding the end of stimulus fixed and moving backwards in time) and a larger proportion of each cohort will enter the market every period. This implies that the per period average of the number of entries driven by extensive margin response increases.

The share of extensive margin response out of total short-term response can more explicitly be calculated as

$$\frac{\sum_{t=0}^{t=T} (F(V_t) - N^c)}{\sum_{t=0}^{t=T} N_t} \quad (2.7)$$

Figure 2.10 shows this share for a wide range of (simulated) stimulus durations, confirming that the share of extensive margin indeed increases monotonically with the duration of tax cut.

## 2.5 Conclusion

Fiscal stimulus programs were widely used to boost activity in various markets during the Great Recession. While it is widely believed that these programs are effective in increasing market activity while in place, evidence on their long-term impact has been mixed. Some studies have shown that such stimulus schemes have considerable permanent effect, while others find seemingly contradictory evidence showing that the boost in market activity created by a stimulus program is reversed shortly after its end and the long-term effect of the stimulus is practically zero.

In this paper, I take a first step towards reconciling these seemingly contradictory results by arguing that the permanent effect of a stimulus is directly related to its duration which could explain why studies that look at stimulus

with considerably different durations arrive at contradicting conclusions. The first part of the paper presented empirical evidence from the UK housing market showing that a temporary tax cut in this market has had a considerable long-term impact on market activity. The second part of the paper highlighted the link between the length of the program and its long-term impact through a search model of the housing market. The intuition captured by this model is that a longer stimulus program gives individuals more time to search, which means they are more likely to be able to find a match before the stimulus ends, which increases the value of entering the market. This in turn implies that a larger proportion of agents will decide to enter the market every period and the share of extensive margin response will grow with the duration of the stimulus.

## 2.6 Figures and Tables

**Table 2.1:** Property transactions tax rates

Price	Average tax rate
up to £125,000	0%
from £125,001 to £250,000	1%
from £250,001 to £500,000	3%
from £500,001 to £1,000,000	4%
from £1,000,001 to £2,000,000	5%
over £2,000,000	7% (bought by individuals) 15% (bought by corporations)

(a) SDLT rates until December 2014

Price	Marginal tax rate
up to £125,000	0%
from £125,001 to £250,000	2%
from £250,001 to £925,000	5%
from £925,001 to £1,500,000	10%
over £1,500,000	12%

(b) SDLT rates since December 2014

Price	Marginal tax rate
up to £145,000	0%
from £145,000 to £250,000	2%
from £250,000 to £325,000	5%
from £325,000 to £750,000	10%
over £750,000	12%

(c) LBTT rates since April 2015

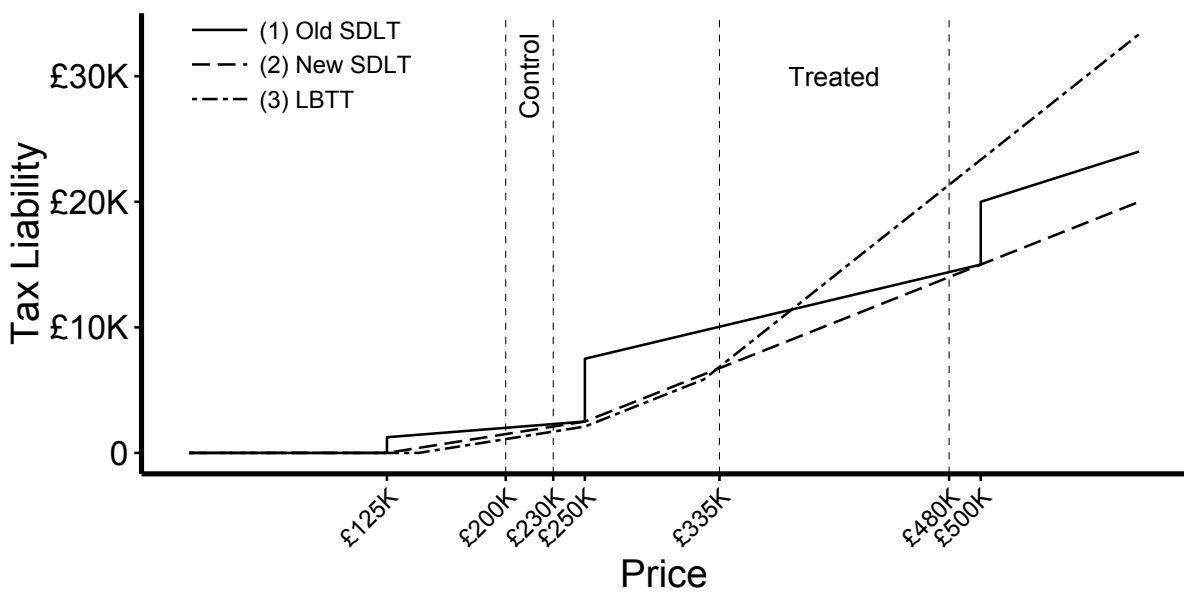
**Notes:** The table shows tax rates under the three different tax schedules in effect in England and Scotland between December 2014 and April 2015.

**Table 2.2:** Difference-in-difference regression results

	<i>Dependent variable:</i>	
	Log number of transactions	
	(1)	(2)
$\beta_s$	0.475 (0.146)	0.496 (0.181)
$\beta_r$	-0.435 (0.124)	-0.392 (0.164)
$\beta_p$	0.058 (0.172)	0.049 (0.099)
Share of extensive margin	0.542 (1.254)	0.605 (1.552)
Month fixed effects	No	Yes
Observations	142	142

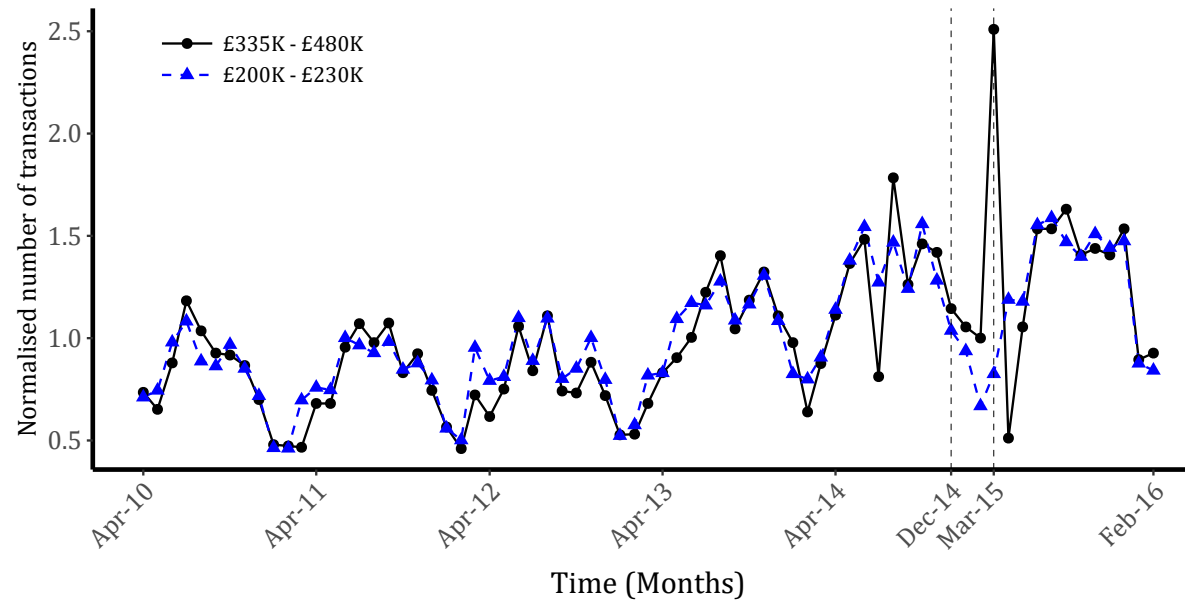
**Notes:** The table reports estimates from equation (2.1).  $\beta_s$  estimates the average percentage increase in number of transactions during the tax cut.  $\beta_r$  estimates the average percentage fall in number of transactions in the reversal period after the end of the tax cut. The reversal period is chosen to be two months. Results are robust to this choice.  $\beta_p$  estimates the difference between treated and control series after the end of reversal.  $\beta_p$  being small and insignificant ensures that the two series have converged and reversal is over. The bottom row reports the share of extensive margin response in total short-term response as  $1 - \frac{2|\beta_r|}{4\beta_s}$ . For regression coefficients HAC standard errors are reported. The standard error for share of extensive margin is calculated using the delta method.

**Figure 2.1:** Property transactions tax liability



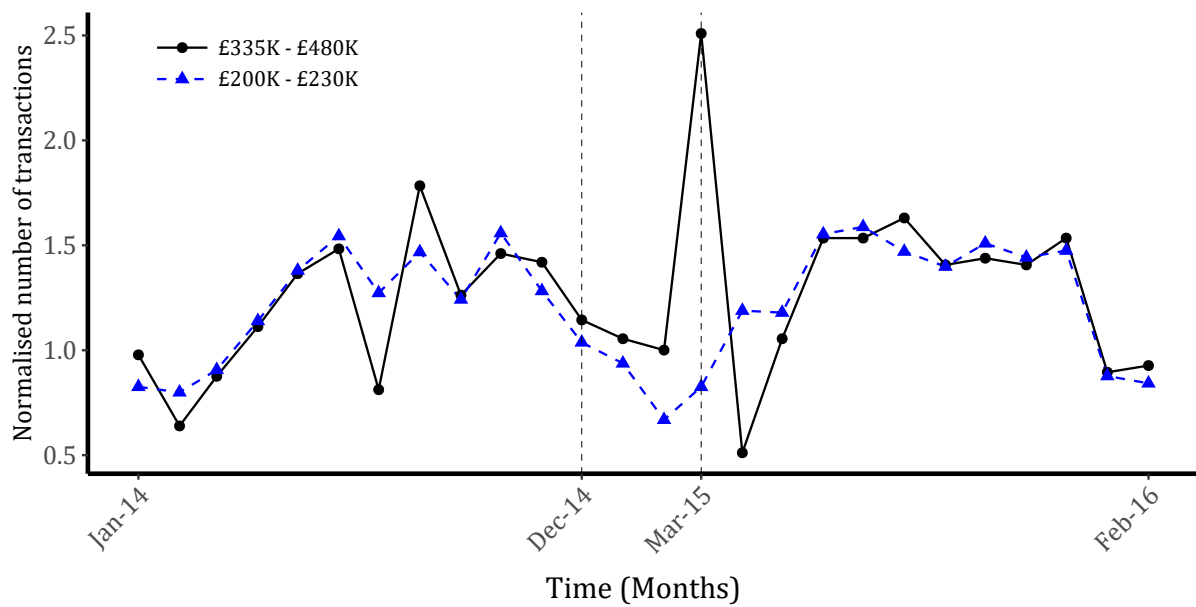
**Notes:** The figure shows tax liability for residential properties in the UK under three different policy regimes. The Old SDLT was in effect in both England and Scotland until December 2014 when New SDLT was announced and immediately implemented. The LBTT took effect in Scotland in April 2015. It was announced in October 2014. The combination of these reforms creates a four-month-long tax cut in the [£335K, £480K] price range, which is used as the treated bracket, where tax liability is reduced unexpectedly in December (from Old SDLT to New SDLT) and is expected to increase in April 2015 (from New SDLT to LBTT).

**Figure 2.2:** Number of transactions in the treated and control price brackets.



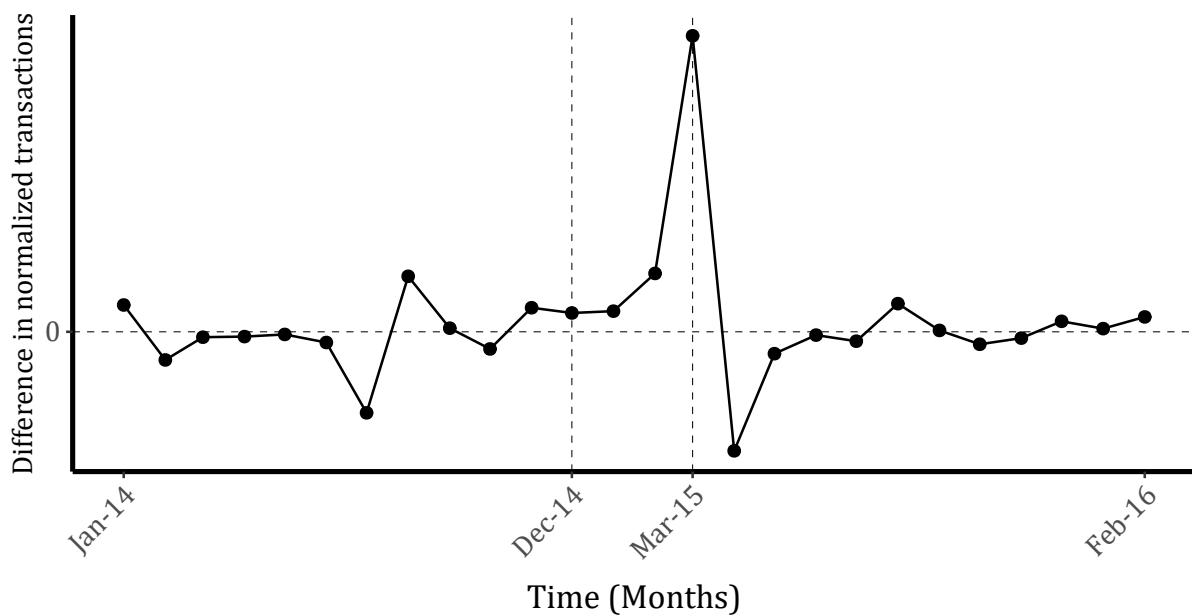
**Notes:** Effect of the temporary tax cut on the number of transactions. Black circles corresponds to the treated price range and blue triangles show the counter-factual from the control price bracket. The two series follow each other very closely until December 2014. They also resume to have parallel trends a few months after April 2015, once the reversal period is over.

**Figure 2.3:** Number of transactions in the treated and control price brackets.



**Notes:** Effect of the temporary tax cut on the number of transactions. Black circles corresponds to the treated price range while blue triangles show the counter-factual.

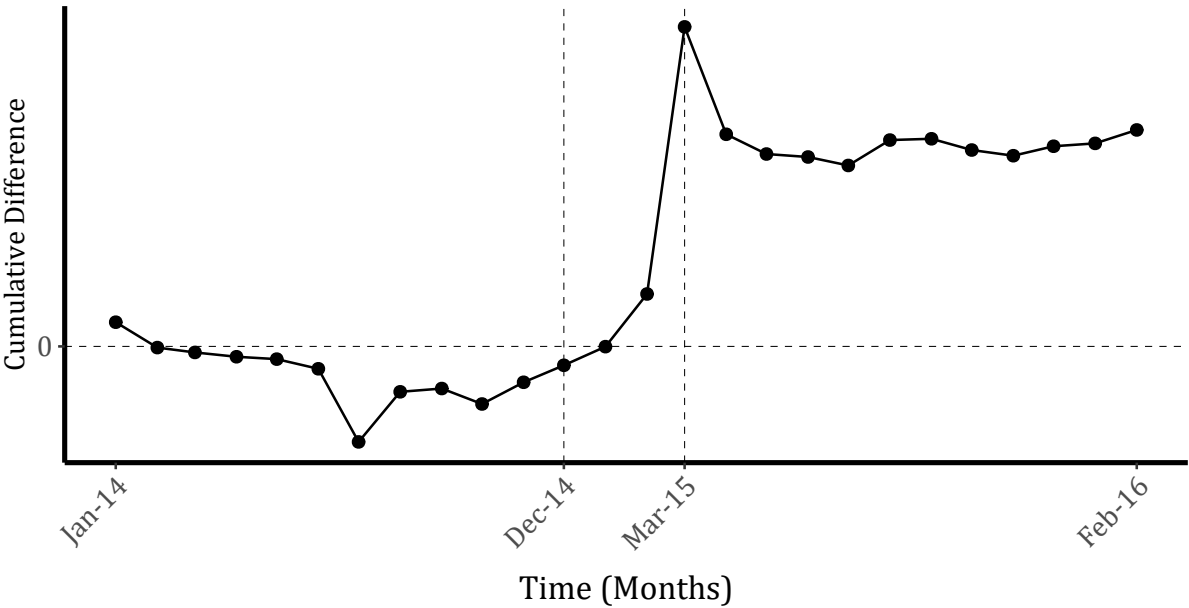
**Figure 2.4:** The difference in activity between the treated and the control price brackets.



**Notes:** The figure shows difference in normalised number of transactions between the treated and the control price brackets. The difference in activity is stable around zero (apart from the run up to the Scottish independence referendum). It increases after the introduction of lower SDLT rates in December 2014 and peaks just before higher LBTT rates were due to take effect in April 2015. Consequently, the difference becomes negative after the end of this period for two months before converging back to zero. This pattern closely matches the prediction of the search model presented in section 2.3.

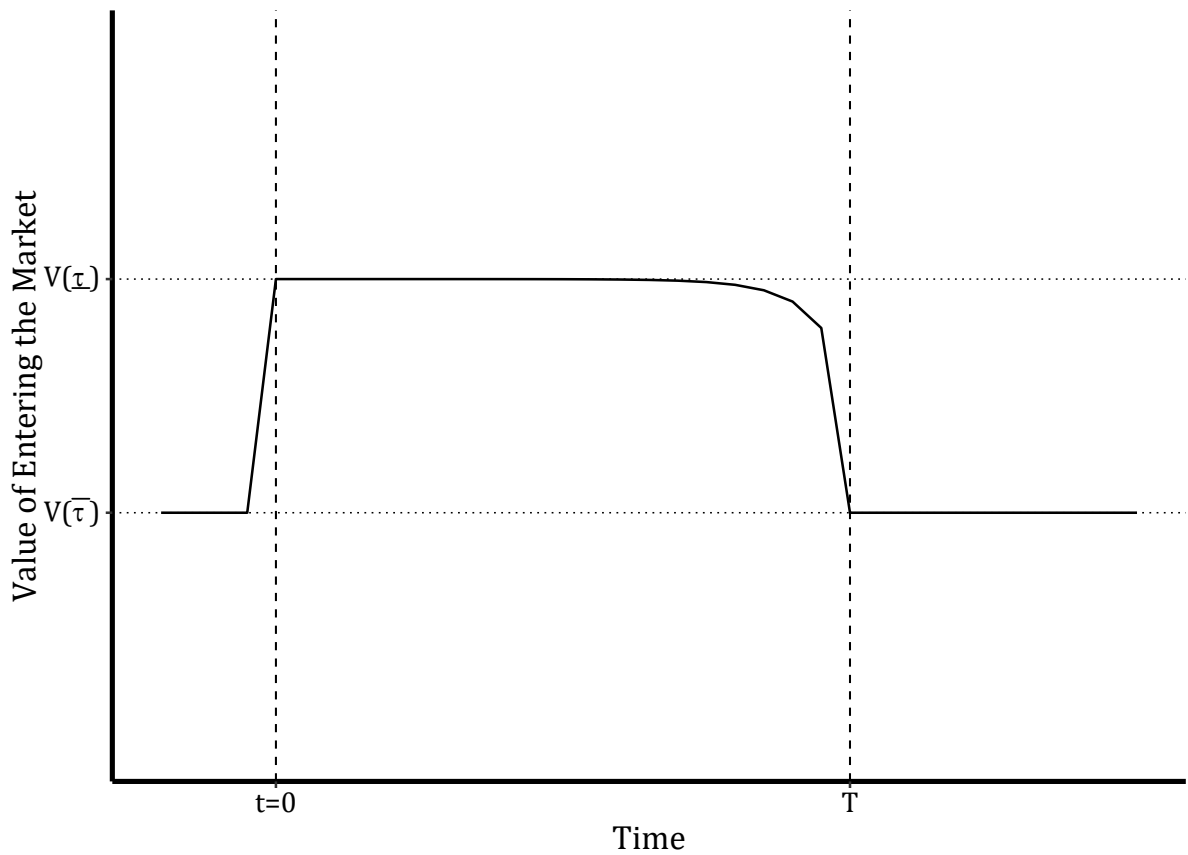


**Figure 2.5:** Cumulative difference in market activity between treated and control price brackets.



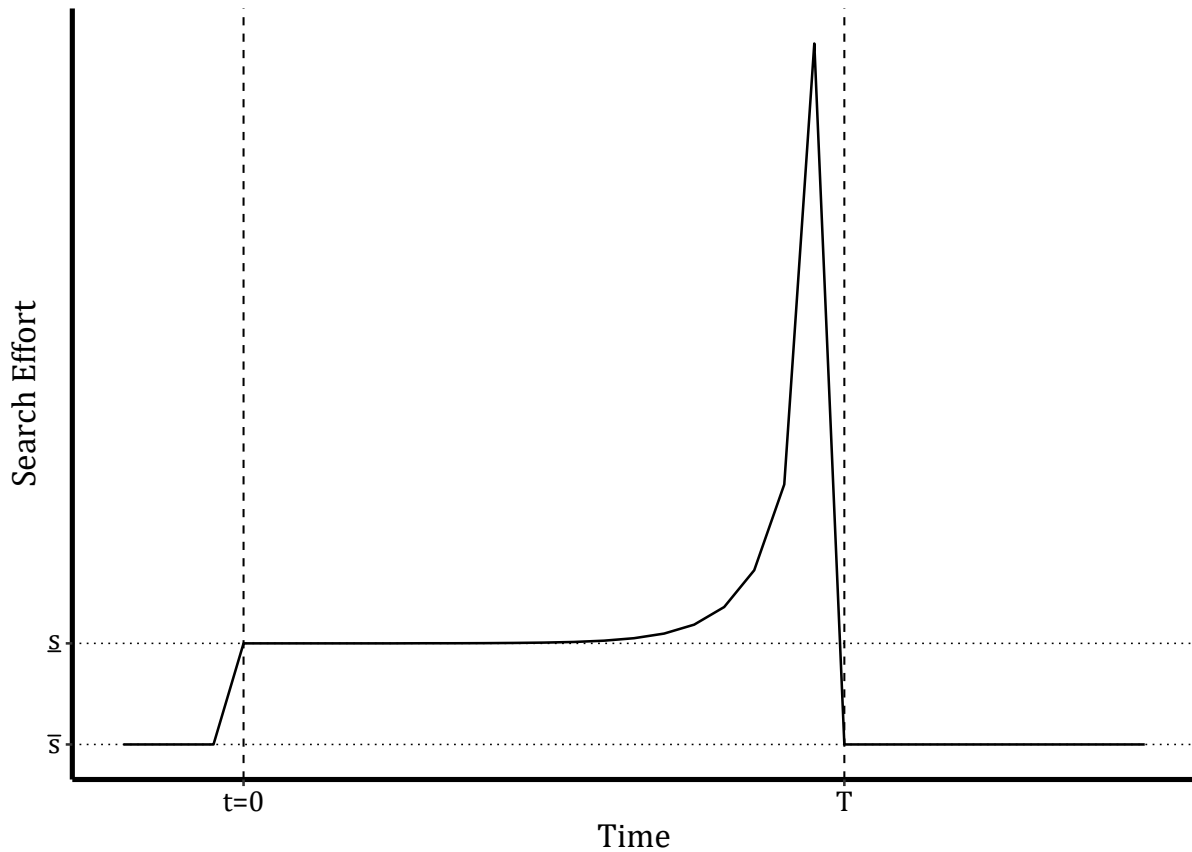
**Notes:** The figure shows the lasting impact of the temporary tax cut on the level of market activity. It shows cumulative difference in normalised number of transactions between the treated and the control price brackets. The cumulative difference rises sharply during the tax cut period and falls consequently. However, it converges to a higher level than before the reforms suggesting that the amount of initial increase in activity was larger than the consecutive reversal and the tax cut. This suggests that a considerable fraction of the response to the tax cut has been along the extensive margin as opposed to time shifting.

**Figure 2.6:** The value of entering the market over time.



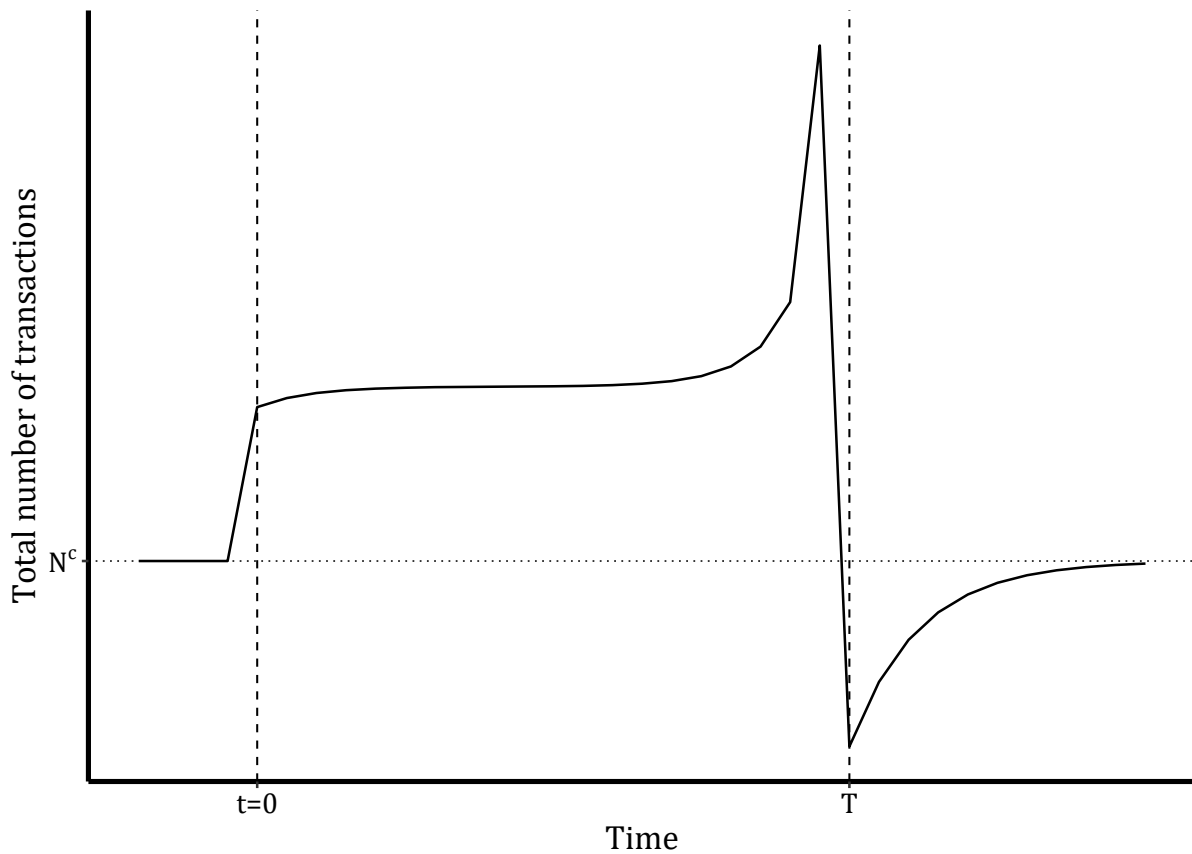
**Notes:** The figure shows the value of entering the market over time. This is constant before and after the temporary tax cut as tax rate is not expected to change and we are in a steady state. During the tax cut, the value of entering the market and starting to search first increases and then gradually declines as we approach the end of the tax cut. This is because the later one enters the market the less likely one is to find a match before the tax rate increases again. This implies that longer tax cuts induce a larger increase in the value of searching and more people will decide to enter the market every period compared to a shorter tax cut.

**Figure 2.7:** Search effort over time.



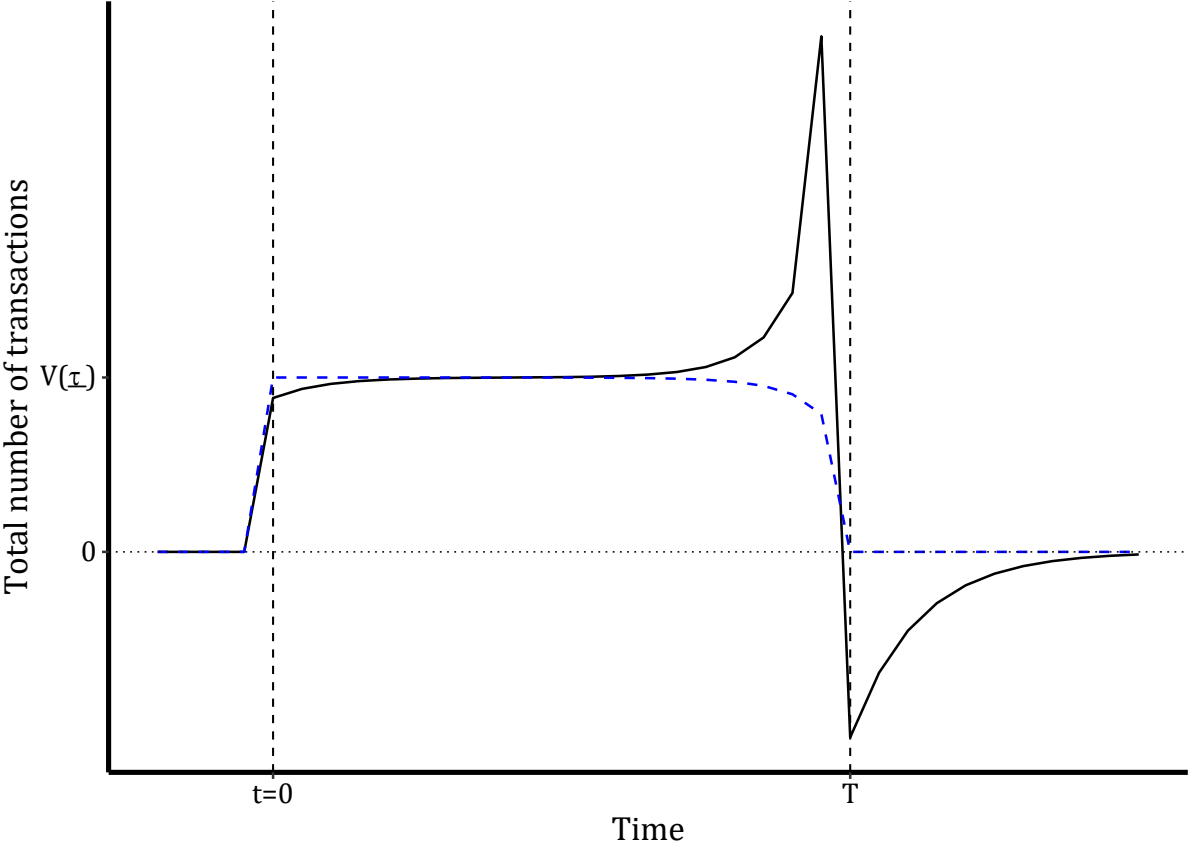
**Notes:** The figure shows the level of search effort, which determines the probability of finding a match, over time. The search effort is constant in the steady states before and after the tax cut. It increases at the beginning of the tax cut as the value of finding a match increases. As we approach the end of the tax cut search effort rises sharply as agents try to find a match before the end. This rise in search effort creates the re-timing response where some transactions that would have happened after  $t = T$  are moved forward in time.

**Figure 2.8:** Simulated number of transactions over time.



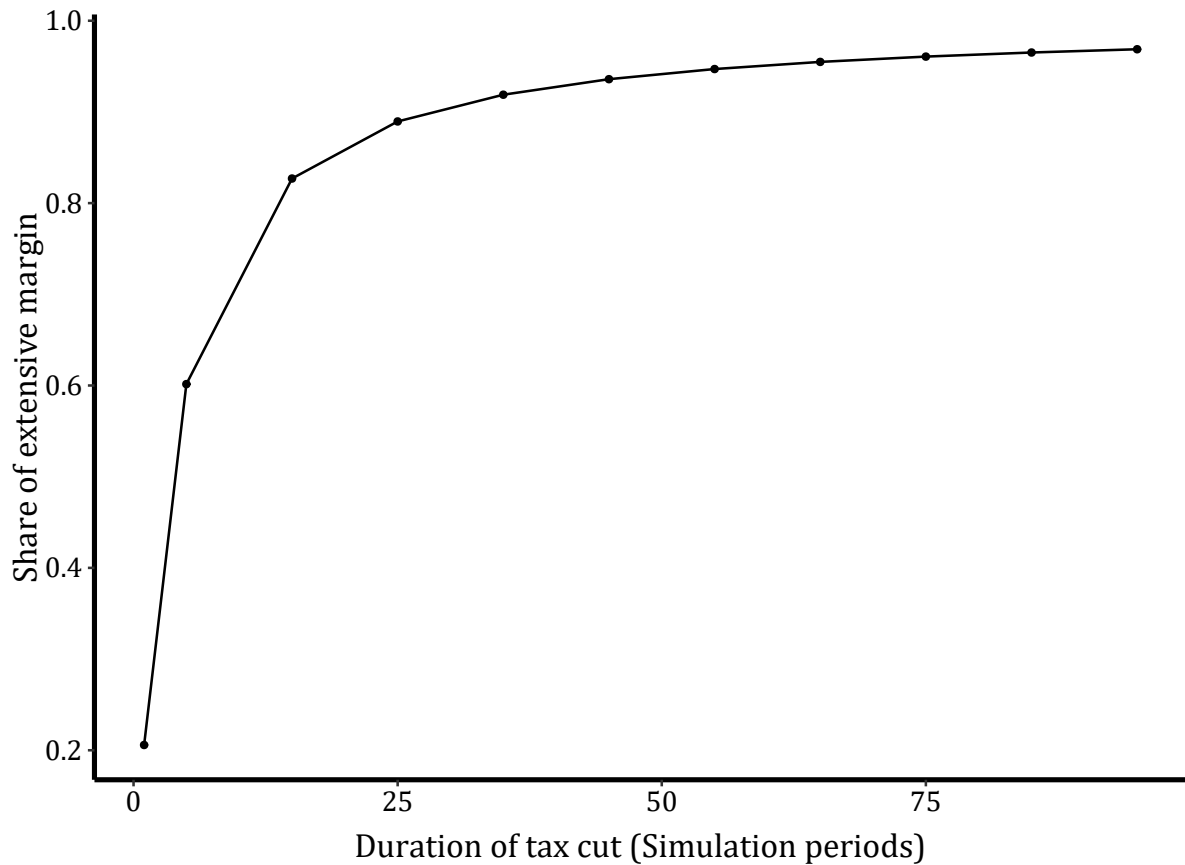
**Notes:** The figure shows the simulated number of matches in the market using Equation (2.2).  $N^c$  marks the high-tax steady state level of activity. The number of matches increases immediately at  $t = 0$  when the tax rate goes down. Towards the end of tax cut, number of matches increases as agents increase their search effort to find a match before the tax rate goes up at  $t = T$ . This implies that an increasing number of agents leave the market just before  $t = T$ . This in turn results in a sharp drop in market activity just after  $t = T$ .

**Figure 2.9:** Number of matches and the number of entries to the market, in excess of high-tax steady state.



**Notes:** The figure shows the simulated number of matches in the market using Equation (2.2) (black solid line) along with number of entries to the market (dashed blue line). The level of activity in the high-tax steady state has been normalised to zero and so the graph shows activity in excess of the high-tax steady state. The area under the black curve measures the short-term response to the tax cut, i.e. the increase in the number of transactions during the tax cut. The area under the new entries curve measures the extensive margin response, i.e. the number of people that enter the market who would not have done so in the absence of the tax cut. The figure illustrates that the response to a short tax cut will be dominated by a large increase in short-term activity while failing to create a large extensive margin response.

**Figure 2.10:** Share of extensive margin response in the total short-term response.



**Notes:** The figure shows the share of extensive margin response out of total short-term response as a function of the duration of the stimulus. The share is calculated using equation (2.7) and simulated tax cuts with various durations. As the duration of the tax cut increases, the value of entering the market in earlier periods increases and more people enter the market every period. This increases the (per period) average of the extensive margin response, which is the permanent component of the effect of the tax cut.

## Chapter 3

# Frequent Repayments as a Screening Device in Micro-lending Under Individual Liability

### **3.1 Introduction**

Microfinance has been of intense academic interest to development economists for the past two decades. Numerous theoretical and empirical studies have been carried out and a vast literature has been produced on the topic. A considerable part of the theoretical literature has been concerned with proposing new methods to improve on existing practices, as well as offering explanations for the unexpected success of microfinance industry in sustainable banking with the poor.

One of the potential difficulties in lending to the poor is the problem of adverse selection. In a credit market with two types of borrowers, risky and safe, where the bank is unable to distinguish between the two types, it has to offer the same interest rate to all borrowers. Then the presence of risky borrowers might drive safe borrowers out of the market, through pushing the break-even interest rate of the bank above their participation threshold (Stiglitz and Weiss 1981). In the context of micro-lending, the main focus of the literature has been on the role of joint liability in tackling the problem of adverse selection. Ghatak (2000) and Van Tassel (1999) for example, address

the problem of adverse selection in the context of microfinance and show how joint liability can be used as a screening device.

But joint liability is not the only form of contract that MFIs use. In fact, a large number of MFIs have abandoned explicit joint liability during the past decade and opted for individual liability loans. In the absence of explicit joint liability, lenders use other techniques, such as non-refinancing threats, sequential lending and frequent repayments, to maintain high repayment rates. The latter especially, is widely believed, at least by practitioners, to be crucial to inducing financial discipline and high repayment rates. This paper seeks to explain the role of frequent repayment in screening different types of borrowers.

Empirical evidence on the effect of repayment frequency is mixed. While some MFIs, such as BRAC, have experienced increases in delinquency rates when trying to reduce the frequency of repayments, there are other studies that find contrasting results. McIntosh (2008) uses variations in the contract terms offered by FINCA Uganda in different parts of the country to investigate the effect of repayment frequency and finds that when groups were allowed to switch to biweekly repayment, both dropout and default rates fell. This experiment investigates the effect of letting existing customers choose from a menu of contracts. Field and Pande (2008) on the other hand, carry out an experiment on groups of first time borrowers and find no significant effect of tighter repayment schedule on default and delinquency rates.

The theoretical literature however, had largely overlooked the role of frequent repayments until recently. Jain and Mansuri (2003) argue that a tight repayment schedule forces clients to borrow from local informal lenders in order to cope with their repayment obligation. This enables the MFI to use the monitoring abilities of informal lenders. Fischer and Ghatak (2010) focus on borrower behavior and argue that frequent repayments could increase the maximum incentive compatible loan size for present biased agents.

This essay also focuses on the behavior of borrowers. However, it assumes that agents are rational in its classic sense, and tries to highlight the role of frequent repayments in screening out risky clients. The basic idea is that a “risky” borrower has a more volatile stream of earnings which makes her



incapable of fitting into a tight repayment schedule. The lender could use this to design a menu of contracts that would separate risky clients from safe ones.

The rest of the paper is organized as follows. Section 3.2 demonstrates the problem of underinvestment in credit markets with adverse selection using a simple model. Section 3.3 expands this baseline model and introduces gradual repayments. Section 3.4 concludes.

## 3.2 One-Shot Repayment

In this section I briefly present a one-period model of adverse selection in credit markets, in which borrowers repay their loan in one single installment, and which closely follows the individual liability lending model by Ghatak (2000). I will then extend the model by introducing the possibility of gradual repayment of the loans, in order to demonstrate the use of frequent repayment as a screening device.

### 3.2.1 The Environment

The economy consists of a lender and a large population of borrowers, the size of which is normalized to one. All agents have access to a risky investment project which needs one unit of capital. Their wealth is assumed to be zero and they have to borrow to finance their projects. Agents have different intrinsic probabilities of success in their investment projects. In particular, there are two types of agents in this economy: a portion,  $\theta$ , of borrowers are “risky”, whose projects succeed with probability  $p_r$ , and the remaining  $1 - \theta$ , are “safe” borrowers who succeed with probability  $p_s$ , with  $0 < p_r < p_s < 1$ . The project yields 0 if it fails and  $R_i > 0$ ,  $i = r, s$ , for a borrower of type  $i$ , if it succeeds. Borrowers are risk neutral and maximize expected returns. The reservation utilities of both types of borrowers are normalized to zero for simplicity.

The lending side consists of a single risk neutral bank whose opportunity cost of capital is  $\rho$  per loan.

The type of a borrower is her private information and the bank cannot observe it. But the outcomes of the borrowers’ projects are observable for the bank. Borrowers have to repay the full cost of the loan when their projects are successful. However, because of the limited liability assumption, which

will be maintained throughout, in the case of failure borrowers are only liable up to the amount of assets they own, which we have assumed to be zero. Enforcement costs are assumed to be negligible for the moment. But we will relax this assumption later.

### 3.2.2 The One-Period Model

If the bank can observe the types of the borrowers, it can charge borrowers of each type a different interest rate such that it breaks even. In particular, I assume that the bank maximizes the utility of borrowers, subject to a zero profit constraint and the limited liability assumption. Solving the zero profit constraint of the bank, we get:

$$r_i = \frac{\rho}{p_i}, \quad i = r, s \quad (3.1)$$

#### 3.2.2.1 The Problem of Underinvestment

Assume that the expected return to all projects are the same,  $p_r R_r = p_s R_s = \bar{R}$  and  $\bar{R} > \rho$ , so that all projects are socially productive. If the bank cannot distinguish between risky and safe borrowers, offering different contracts is not possible, as risky borrowers would have an incentive to pretend to be safe and pay the lower interest rate  $\rho/p_s$ . So the bank has to offer a single interest rate to both types of borrowers through which it can break even.

If the bank charges everyone the same interest rate, safe borrowers will have a lower expected payoff, as they are more likely to succeed and have to repay more often. But the expected return of a safe borrower is the same as that of a risky borrower. Now solving the zero profit constraint of the bank, we get  $r = \rho/\bar{p}$ , where  $\bar{p} = \theta p_r + (1 - \theta)p_s$ . Denote by  $U_i(r)$  the utility of a type  $i$  borrower as a function of the interest rate. Since  $U_s(r) < U_r(r)$ , for  $r = \rho/\bar{p}$  to be the optimal pooling equilibrium, we need to make sure that safe borrowers will participate under this interest rate, i.e.,  $\rho/\bar{p} \leq \bar{R}/p_s$

However, if the above inequality fails to hold, a pooling contract that would attract both types of borrowers does not exist. The optimal contract then, would be the one that only attracts risky borrowers, charges an interest of  $\rho/p_r$  and satisfies the zero-profit condition for lending to risky borrowers

only. In this case, safe borrowers will be forced out of the market by the high interest rate and only risky borrowers will borrow. Although the projects of safe borrowers are socially productive, they will not invest and both the repayment rate and welfare are strictly lower than the full-information case. This is known as the under-investment problem in credit markets with adverse selection (Stiglitz and Weiss 1981).

### 3.3 Frequent Repayment as a Screening Device

This section develops a model of gradual repayments and shows how it can be used to design a menu of separating contracts for two types of borrowers. The basic idea here is as follows. Since a risky borrower succeeds less often, but earns bigger sums when she is successful, the variance of her earnings is bigger than that of a safe borrower. A tight repayment schedule, i.e., higher repayment frequency, requires the borrowers to have a steady stream of returns, as they are required to repay more often. The lender can use this feature to make the borrowers self-select into two different contracts. Those who have a more volatile flow of income cannot fit into a tight repayment schedule, and would rather pay a higher interest rate in return for having a more flexible repayment schedule, i.e., repaying less often.

Since we are going to make a model of gradual repayment, we need to introduce a process for the earnings of the borrowers which would specify how they earn money over the course of the loan.

#### 3.3.1 Returns of Investment Projects

As the first departure from the baseline model, we need to decide how the returns of the projects are realized over time. I assume that the (accumulated) returns of an investment project constitute a Poisson process. In particular, borrowers of type  $i$  earn money at random points in time, and at a constant rate of  $\lambda_i$ ,  $i = r, s$ , with  $\lambda_r < \lambda_s$ . The probability of  $k$  successes per unit of time is given by the Poisson probability distribution function:

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (3.2)$$

Each time a type  $i$  borrower is successful she earns a fixed amount,  $R_i$ . So, the expected return from the projects are  $\lambda_i R_i$ <sup>1</sup>. This is assumed to be the same for both types of borrowers as in the one period model, i.e.,  $\lambda_r R_r = \lambda_s R_s = \bar{R}$ . Since  $\lambda_r < \lambda_s$ , we have  $R_r > R_s$ .

One way to interpret this is to assume that each borrower of type  $i$  has a shop and customers arrive at her shop and buy products at a rate  $\lambda_i$ . Risky borrowers earn higher profits per sale, but manage to sell their products less often. Safe borrowers on the other hand earn lower profits per sale, but sell more often. Although the expected returns of both borrowers are equal to  $\bar{R}$ , the variance of the profits of the safe borrower is smaller.

### 3.3.1.1 Equivalence to the one-shot repayment model

The model with a Poisson earnings process can be shown to be equivalent to the one-shot repayment model of section 3.2 provided that the following assumption holds; we need to assume that the profits earned from one success, which is equal to  $R_i$  for a borrower of type  $i$ , is sufficient to repay the loan, that is:

**Assumption 3.1**  $R_i \geq r_i$ .

In this case the probability that a borrower of type  $i$  will be able to repay the loan is equal to the probability that she is successful at least once. Using the distribution function of the Poisson distribution, this implies:

$$p_i = 1 - e^{-\lambda_i} \quad (3.3)$$

where, as in the one-shot repayment model,  $p_i$  is the probability that a type  $i$  borrower will be able to repay her loan.

---

<sup>1</sup>The time period over which the returns of projects are realized is normalized to 1 and the mean of the Poisson distribution is equal to the rate of arrivals, in this case,  $\lambda$ .

### 3.3.2 Preferences over Contracts and the Single-Crossing Property

#### 3.3.2.1 Gradual Repayment and the Timing of Installments

Contracts consist of a pair  $(n, r)$ , where  $r$  is the interest rate and  $n$  is the number of installments. I assume that the loan is disbursed at  $t = 0$ , the project starts immediately and the repayment ends at  $t = 1$ . So, a borrower who has accepted a contract  $(n, r)$  has to repay  $n$  installments at  $t = k/n$ , for  $k = 1, \dots, n$ , each of size  $r/n$ . I assume that borrowers do not discount future payoffs.

Moreover, I assume that only the profits realized between  $t = (k - 1)/n$  and  $t = k/n$  can be used to pay the  $k$ -th installment. In other words, borrowers cannot save between periods. So, a borrower can repay her  $k$ -th installment only if she has been successful at least once between  $t = (k - 1)/n$  and  $t = k/n$ .

**Assumption 3.2** *Borrowers do not save between periods.*

#### 3.3.2.2 Repayment Incentive Constraint: Avoiding Strategic Default

In this section I relax the assumption that enforcement costs are negligible and assume that the lender is not able to enforce contracts. This means that the lender has to design the contracts such that the borrower would prefer repayment over strategic default.

I assume that every time a borrower defaults, i.e., is unable or unwilling to repay an installment, she incurs a constant disutility or cost of default,  $c$ . An alternative would be to assume that the cost of default is proportional to the size of the missed payment. While this could be a more realistic assumption, assuming a fixed cost makes the model more tractable and could be a reasonable proxy for the cost of default. One plausible interpretation for this would be the increase in the probability of losing the next loan. For example, in the Grameen II system of the Grameen Bank, when a borrower defaults on her loan, “her loan ceiling that she has built over years, gets wiped out” (Yunus 2002). For the sake of tractability, I will also assume that borrowers

cannot pay previously missed instalments.

Therefore, I focus my attention on the set of contracts that provide the borrower with enough incentive to avoid strategic default, namely, contracts for which  $r/n < c$ . If this does not hold, borrowers would prefer to default on every installment, even when they are able to repay.

### 3.3.2.3 Preferences Over Contracts

Provided that assumptions 3.1 and 3.2 hold, the expected utility of a borrower of type  $i$  from accepting a contract  $(n, r)$  is:

$$U_i(n, r) = \bar{R} - \frac{r}{n}(1 - e^{-i/n})n - nce^{-i/n} \quad (3.4)$$

Remember that  $1 - e^{-i/n}$  is the probability that a borrower can afford to repay an installment. So, the second term is the expected number of installments that the borrower is going to pay, times the size of each installment,  $r/n$ . The third term is the expected number of defaults times the cost of each default. This can be rewritten as:

$$U_i(n, r) = (\bar{R} - r) + (r - nc)e^{-i/n} \quad (3.5)$$

For simplicity I will treat  $n$  as a continuous variable henceforth. This will simplify the model without changing the qualitative implications.

**Lemma 3.1** *The preferences of borrowers over contracts satisfy the single-crossing property.*

Proof. See the Appendix A.

The single-crossing property is a standard feature required for incentive compatibility in screening different types of agents in adverse selection models.

Formally, the single-crossing property means that the slope of the indifference curve of a borrower is an increasing function of her type,  $\lambda$  in this case (see Figure 3.1), which means that the marginal rate of substitution between  $r$  and  $n$  is different for the two types of borrowers. In particular, a rise in the interest rate has a smaller effect on the expected utility of a risky borrower, as she repays less often. In contrast, she cares about the number of installments

more than a safe borrower does, because the risky borrower is more likely to default and incur the cost,  $c$ . So, the risky borrower is willing to accept a bigger rise in the interest rate, in return for a smaller decrease in the number of installments, compared to the safe borrower. This makes the indifference curve of the risky borrower steeper than that of the safe borrower.

### 3.3.3 Optimal Contracting with Gradual Repayment

This section characterizes a menu of optimal contracts. This is done by solving for optimal separating contracts in a contracting problem with the following timing:

1. The lender offers the borrower a menu of contracts,  $\{(n_r, r_r), (n_s, r_s)\}$ .
2. The borrower decides whether to participate and chooses a contract.
3. Loans are disbursed and the borrower starts her investment project.
4. For each  $k = 1, \dots, n$ , the borrower decides whether to repay an installment of size  $r/n$ , at  $t = k/n$ , if she can afford it.

The optimal contracting problem is as follows: the lender maximizes a weighted average of welfares of the two types of borrowers by choosing two contracts  $(n_r, r_r)$  and  $(n_s, r_s)$ :

$$W = \gamma U_r(n_r, r_r) + (1 - \gamma) U_s(n_s, r_s) \quad (3.6)$$

where  $\gamma \in (0, 1)$ .

The optimization is subject to the following constraints:

- a) The zero-profit constraint of the lender: The expected profit of the lender from each contract has to be non-negative. The expected number of installments that a borrower of type  $i$  can afford to repay is  $n(1 - e^{-i/n})$ . The size of each installment is  $r/n$ . The zero-profit condition requires the lender to charge the following interest rate:

$$r_i = \frac{\rho}{1 - e^{-i/n}} \quad (3.7)$$

Since the lender is maximizing the welfare of the borrower, the zero-profit condition binds at the optimum. Let  $ZPC_j$  denote the set of contracts that satisfy this constraint with equality.

- b) The Participation Constraint of the borrower ( $PC$ ): The expected payoff of the borrower from accepting the contract has to be as much as her reservation utility, which is normalized to zero. This implies:

$$U_j(n, r) = (\bar{R} - r) + e^{-i/n}(r - nc) = 0 \quad (3.8)$$

- c) The limited liability constraint: A borrower cannot repay more than the amount of assets she owns, which is assumed to be zero here. This constraint along with assumptions 3.1 and 3.2, imply that a borrower will repay the  $k$ -th installment only if she has had at least one success during the last period, that is, between  $t = (k - 1)/n$  and  $t = k/n$ .

- d) The incentive compatibility constraint: The contracts have to be such that a borrower of type  $i$  would prefer the contract meant for her type to the other option. More formally:

$$U_r(n_r, r_r) \geq U_r(n_s, r_s) \quad (3.9)$$

$$U_s(n_s, r_s) \geq U_s(n_r, r_r) \quad (3.10)$$

- e) The repayment incentive constraint ( $RIC$ ): Since contracts are not enforceable, the interest rate and the frequency of installments have to be such that the borrower prefers repayment over strategic default. As discussed earlier, this requires  $r/n \geq c$ .

Before proceeding to the solution of the optimal contracting problem we need to make sure that the set of contracts that satisfy these constraints (except for incentive compatibility), is not empty. Specifically, I assume that single repayment lending to risky borrowers is feasible, i.e., the contract  $(1, \rho/(1 - e^{-r}))$  which has the zero-profit interest rate and only one installment satisfies the participation and repayment incentive constraints.



More formally, for this contract to satisfy the repayment incentive constraint we need:

**Assumption 3.3**  $r_r^1 \geq \frac{c}{1 - e^{-\lambda r}}$

Also, for this contract to satisfy the participation constraint we need to assume that:

**Assumption 3.4**  $\bar{R} - r_r^1 + e^{-r}(r_r^1 - c) \geq 0$

or in terms of the baseline model, with  $p_r = 1 - e^{-r}$  for  $n = 1$ , this can be rewritten as:

$$\bar{R} - r_r^1 p_r \geq c(1 - p_r) \tag{3.11}$$

The left hand side is the expected profit of the investment project and the right hand side is the expected cost of default. Assumption 3.4 says that the expected net gain of investment is such that the borrower chooses to participate.

As Figure 3.2 shows, if the participation constraint of the risky borrower holds, so will the  $PC$  of the safe borrower. Contracts below the  $RIC$  line satisfy the repayment incentive constraint. Therefore, we are interested in contracts that lie on  $ZPC_i$ ,  $i = r, s$ , and are below  $RIC$  and  $PC_r$ .

**Proposition 3.1** *Under assumptions 3.1 - 3.4, optimal separating contracts  $(n_r, r_r)$  and  $(n_s, r_s)$  exist. Furthermore,  $r_s < r_r$  and  $n_s > n_r$ . The welfare of the risky borrower is the same as in the full-information case while the welfare of the safe borrower is strictly lower.*

Proof. See Appendix A.

Figure 3.3 depicts the menu of optimal contracts. A formal proof of the proposition is provided in the appendix. Intuitively, the pair of contracts can be constructed as follows. First, pick the contract  $(n_r, r_r)$  which gives the highest feasible utility to risky borrowers, represented by  $u_r^1$ . Next, pick a contract,  $(n_s, r_s)$ , for the safe borrower that would satisfy the incentive compatibility constraint (along with other constraints). For incentive compatibility, this contract needs to be on or above  $u_r^1$  so that  $u_r(n_s, r_s) < u_r(n_r, r_r)$ , and below  $u_s^1$ , so that  $u_s(n_s, r_s) > u_s(n_r, r_r)$ . Given that the objective is to maximize a

linear combination of utilities, the optimal contract for the safe type is the one that provides the highest feasible utility,  $u_1^s$ .

Proposition 3.1 shows that risky borrowers will choose a contract with a more flexible repayment schedule, i.e. lower frequency of repayment, but a higher interest rate. Safe borrowers however are able to fit in a tight repayment schedule, namely, repaying more frequently, and will choose a contract with a lower interest rate, but higher frequency of repayment. This is in fact a signal that a safe borrower can send to the lender to credibly reveal her type, and one that a risky borrower cannot afford.

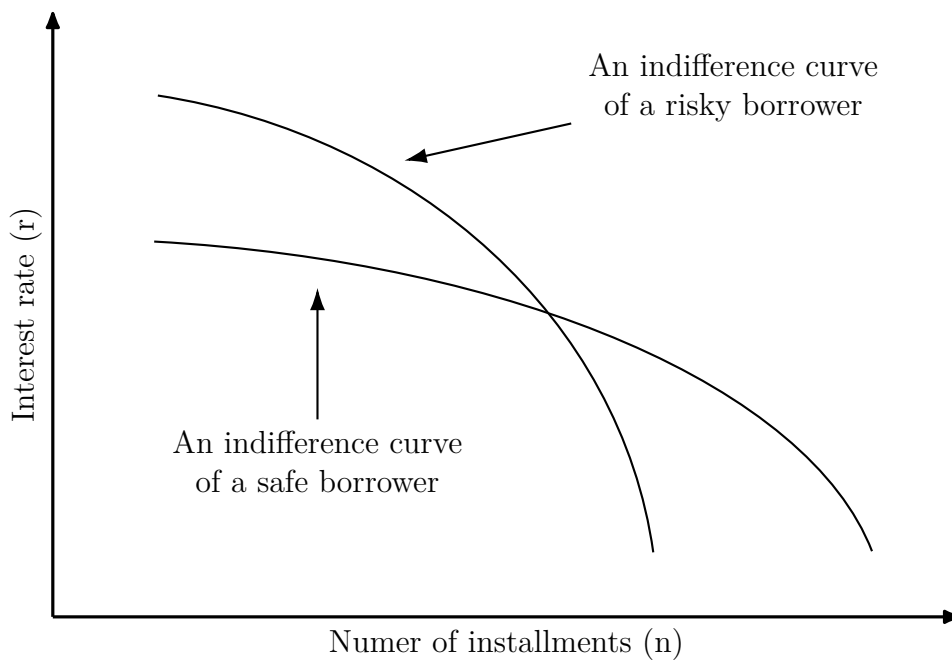
### **3.4 Conclusion**

This essay looks at an economic environment in which inefficiencies are present due to asymmetric information. It was shown using a simple model of adverse selection that under individual liability contracts, a tight repayment schedule can be used to screen different types of borrowers. This is done by offering the clients a menu of contracts which is designed such that borrowers of different types would self-select into different contracts.

This is in line with the almost universal belief among microfinance practitioners that a frequent repayment schedule is crucial to maintaining high repayment rates and avoiding default. This model should be seen as a first attempt to suggest an explanation for the effect of frequent repayment.

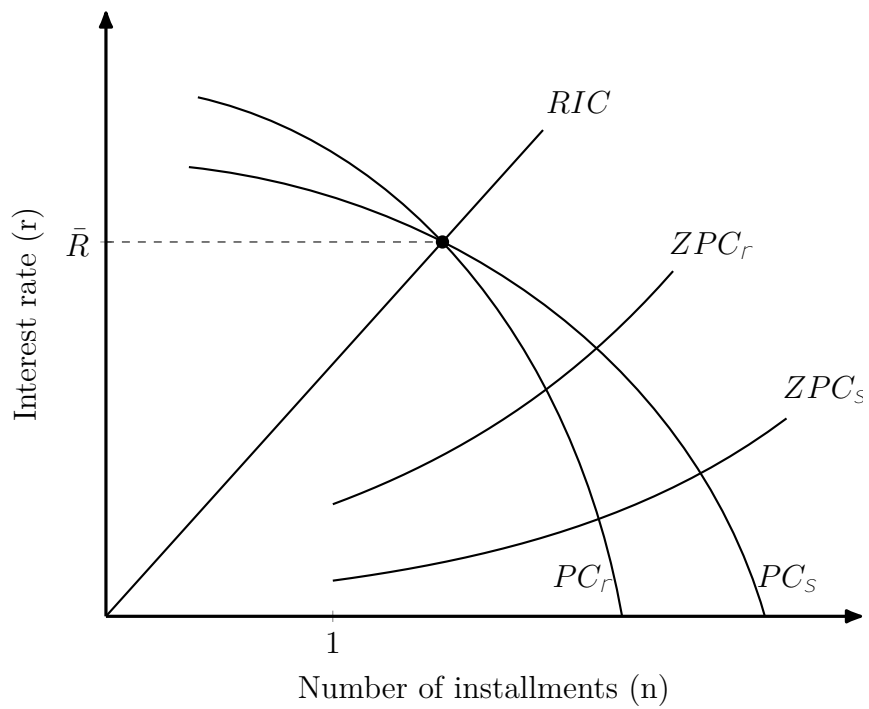
### 3.5 Figures

**Figure 3.1:** Indifference curves of risky and safe borrowers



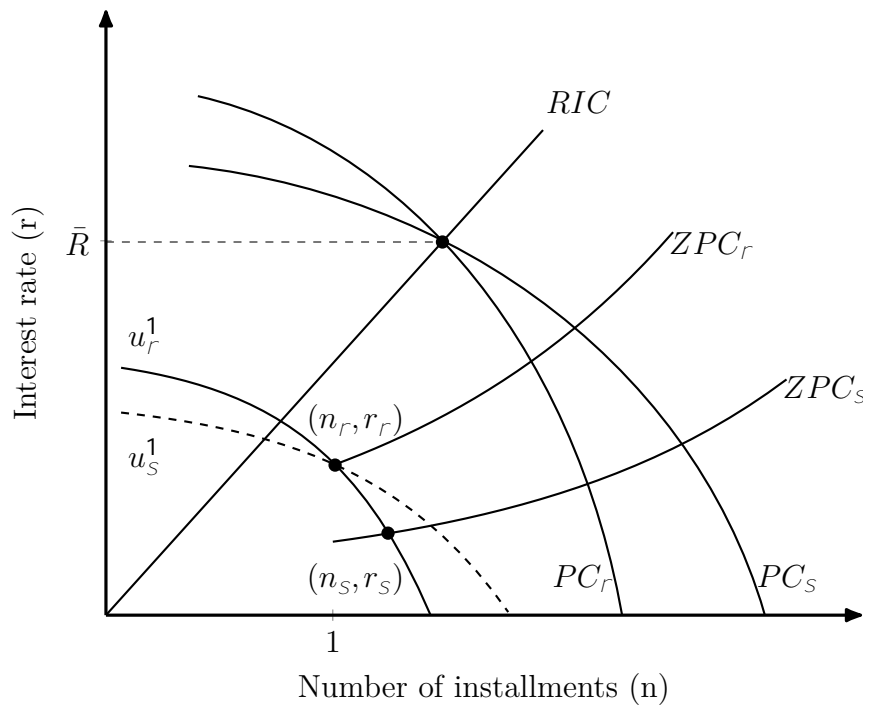
**Notes:** The figure shows an indifference curve of a risky and a safe borrower. The marginal rate of substitution between  $n$  and  $r$  is an increasing function of the type of agents,  $\lambda$ . Therefore, the indifference curves satisfy the single-crossing property. Note that utility increases towards the origin.

**Figure 3.2:** The constraints of the contracting problem



**Notes:** The figure shows the constraints of the lender's optimization problem. Contracts below  $RIC$ ,  $PC_s$  and  $PC_r$  curves satisfy the corresponding constraints. Contracts that lie on  $ZPC_s$  and  $ZPC_r$  satisfy the zero profit constraint for safe and risky borrowers, respectively.

**Figure 3.3:** Optimal separating contracts



**Notes:** The figure shows the optimal separating contracts  $(n_s, r_s)$  and  $(n_r, r_r)$ . These contracts are below  $RIC$ ,  $PC_s$  and  $PC_r$  and therefore satisfy these constraints. They also lie on their corresponding  $ZPC$  curve which means they satisfy the zero-profit condition. They also satisfy the incentive compatibility constraint as the risky type is indifferent between the two contracts, while the safe borrowers strictly prefers  $(n_s, r_s)$ .

# References

- Agarwal, Sumit, Liu Chunlin, and Nicholas S Souleles. 2007. “The Reaction of Consumer Spending and Debt to Tax Rebates-Evidence from Consumer Credit Data”. *Journal of political economy* 115 (6): 986–1019.
- Auerbach, Alan J, and James Poterba. 1988. “Capital gains taxation in the United States: Realizations, revenue, and rhetoric”. *Brookings Papers on Economic Activity* 1988 (2): 595–637.
- Baily, Martin Neil. 1978. “Some aspects of optimal unemployment insurance”. *Journal of Public Economics* 10 (3): 379–402.
- Best, Michael Carlos, and Henrik Jacobsen Kleven. 2017. “Housing market responses to transaction taxes: Evidence from notches and stimulus in the UK”. *The Review of Economic Studies* 85 (1): 157–193.
- Blau, David M, and Philip K Robins. 1986. “Job search, wage offers, and unemployment insurance”. *Journal of Public Economics* 29 (2): 173–197.
- Burman, Leonard E, and William C Randolph. 1994. “Measuring Permanent Responses to Capital-Gains Tax Changes in Panel Data”. *The American Economic Review*: 794–809.
- Chetty, Raj. 2006. “A General Formula for the Optimal Level of Social Insurance”. *Journal of Public Economics* 90 (10): 1879–1901.
- Chetty, Raj, John N Friedman, and Emmanuel Saez. 2013. “Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings”. *The American Economic Review* 103 (7): 2683–2721.

- Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence". *The American Economic Review* 99 (4): 1145–1177.
- Field, Erica, and Rohini Pande. 2008. "Repayment frequency and default in microfinance: evidence from India". *Journal of the European Economic Association* 6 (2-3): 501–509.
- Fischer, G, and M Ghatak. 2010. "Repayment Frequency and Lending-Contracts with Present-Biased Borrowers". *London School of Economics Working Paper*.
- Fudenberg, Drew, and Jean Tirole. 1991. *Game theory*.
- Ghatak, Maitreesh. 2000. "Screening by the company you keep: Joint liability lending and the peer selection effect". *The Economic Journal* 110 (465): 601–631.
- Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation". *The Journal of Political Economy* 108 (2): 352–378.
- Gruber, Jonathan. 1997. "The Consumption Smoothing Benefits of Unemployment Insurance". *The American Economic Review* 87 (1): 192–205.
- Hopenhayn, Hugo A, and Juan Pablo Nicolini. 1997. "Optimal Unemployment Insurance". *Journal of Political Economy* 105 (2): 412–438.
- Jain, Sanjay, and Ghazala Mansuri. 2003. "A little at a time: the use of regularly scheduled repayments in microfinance programs". *Journal of development Economics* 72 (1): 253–279.
- Johnson, David S, Jonathan A Parker, and Nicholas S Souleles. 2006. "Household Expenditure and the Income Tax Rebates of 2001". *The American Economic Review* 96 (5): 1589–1610.

- Kolsrud, Jonas, et al. 2018. “The optimal timing of unemployment benefits: Theory and evidence from Sweden”. *American Economic Review* 108 (4-5): 985–1033.
- Lalive, Rafael. 2008. “How do extended benefits affect unemployment duration? A regression discontinuity approach”. *Journal of Econometrics* 142 (2): 785–806.
- Lalive, Rafael, Jan Van Ours, and Josef Zweimüller. 2006. “How changes in financial incentives affect the duration of unemployment”. *The Review of Economic Studies* 73 (4): 1009–1038.
- Machin, Stephen, and Alan Manning. 1999. “The causes and consequences of long term unemployment in Europe”. *Handbook of Labor Economics* 3:3085–3139.
- McIntosh, Craig. 2008. “Estimating treatment effects from spatial policy experiments: an application to Ugandan microfinance”. *The Review of Economics and Statistics* 90 (1): 15–28.
- Mian, Atif, and Amir Sufi. 2012. “The Effects of Fiscal Stimulus: Evidence From The 2009 Cash For Clunkers Program.” *Quarterly Journal of Economics* 127 (3).
- Parker, Jonathan A, et al. 2013. “Consumer Spending and the Economic Stimulus Payments of 2008”. *The American Economic Review* 103 (6): 2530–2553.
- Rothstein, Jesse. 2011. *Unemployment insurance and job search in the Great Recession*. Tech. rep. 17534. National Bureau of Economic Research.
- Schmieder, Johannes F, and Till Von Wachter. 2016. “The effects of unemployment insurance benefits: New evidence and interpretation”. *Annual Review of Economics* 8:547–581.



- Schmieder, Johannes F, and Till von Wachter. 2017. "A Context-Robust Measure of the Disincentive Cost of Unemployment Insurance". *American Economic Review* 107 (5): 343–48.
- Shavell, Steven, and Laurence Weiss. 1979. "The Optimal Payment of Unemployment Insurance Benefits Over Time". *Journal of Political Economy* 87 (6): 1347–1362.
- Stiglitz, Joseph E, and Andrew Weiss. 1981. "Credit rationing in markets with imperfect information". *The American economic review* 71 (3): 393–410.
- Van den Berg, Gerard J. 1990. "Nonstationarity in job search theory". *The Review of Economic Studies* 57 (2): 255–277.
- Van Tassel, Eric. 1999. "Group lending under asymmetric information". *Journal of Development Economics* 60 (1): 3–25.
- Werning, Iván. 2002. "Optimal unemployment insurance with unobservable savings". *University of Chicago and UTDT* 4.
- Wolpin, Kenneth I. 1987. "Estimating a structural search model: The transition from school to work". *Econometrica: Journal of the Econometric Society*: 801–817.
- Yunus, Muhammad. 2002. "Grameen Bank II: designed to open new possibilities". *Grameen Dialogue* 50.

## Appendix A

### Proofs for Chapter 3

**Proof of Lemma 3.1.** A sufficient condition for the preferences to satisfy the single-crossing property is<sup>1</sup>:

$$\frac{\partial}{\partial \lambda_i} \left( -\frac{\partial U_i / \partial n}{\partial U_i / \partial r} \right) > 0. \quad (\text{A.1})$$

That is, the slopes of the indifference curves of agents are required to be a monotone function of their types.

The slope of the indifference curve of a type  $i$  borrower is equal to:

$$\frac{dr}{dn} = -\frac{\partial U_i / \partial n}{\partial U_i / \partial r} = \frac{e^{-r/n}}{1 - e^{-r/n}} \left\{ \frac{\lambda}{n^2} (r - nc) - c \right\} \quad (\text{A.2})$$

Differentiating this with respect to  $\lambda$  and simplifying yields:

$$\frac{\partial}{\partial \lambda_i} \left( -\frac{\partial U_i / \partial n}{\partial U_i / \partial r} \right) = \frac{e^{-r/n}}{n^2 (1 - e^{-r/n})^2} \left( r - \frac{\lambda}{n} (r - nc) \right) \quad (\text{A.3})$$

which is positive, given that we are dealing with contracts that satisfy repayment incentive constraint, i.e.  $r < nc$ .

#### **Proof of Proposition 3.1.**

Take  $(1, \rho / (1 - e^{-r}))$  as the contract for risky borrowers and let us denote this contract by  $(n_r, r_r)$ . Assumptions 3.3 and 3.4 guarantee that  $(n_r, r_r)$  satisfies all of the constraints of the contracting problem.

Figure 3.3 shows the constraints of the problem and indifference curves of

---

<sup>1</sup>See Fudenberg and Tirole (1991), chapter 7, p. 259 for a formal definition.

borrowers. Let us denote by  $u_r^1$  the indifference curve of a risky borrower that passes through this contract. Let us also denote by  $u_r^0$  the indifference curve of a risky borrower that represents utility level zero, which coincides with  $PC_r$ .

Since  $u_r^1$  represents a level of utility higher than zero, and utility increases as one moves towards the origin, the  $u_r^1$  curve will always be below  $u_r^0$ . In particular, consider  $n_0$ , the value of  $n$  for which  $u_r^0$  ( $PC_r$ ) intersects  $ZPC_S$ .  $u_r^1$  will be below  $u_r^0$  and therefore below  $ZPC_S$  for this  $n$ . Therefore,  $u_r^1 < ZPC_S$  at  $n = n_0$ .

We know that  $ZPC_S$  is always below  $ZPC_r$ . In particular, at  $n = 1$ , where  $u_r^1$  (by construction) intersects  $ZPC_r$ , it will be above  $ZPC_S$ . So,  $u_r^1 > ZPC_S$  at  $n = 1$ .

It follows from the continuity of  $u_r^1$  and  $ZPC_S$  that they must intersect for some  $n$  between  $n = 1$  and  $n = n_0$ . Let us denote the contract at this point of intersection by  $(n_S, r_S)$ .

Notice that  $n_S > 1 = n_r$ . Equation (A.2) shows that  $\frac{dr}{dn} < 0$  for all points below  $RIC$ . So,  $n_S > n_r$  implies that  $r_S < r_r$ .

It is obvious that  $(n_S, r_S)$  satisfies zero-profit and repayment incentive constraints. It is below  $u_r^0$  and therefore satisfies the participation of the risky borrower. Since the participation constraint of the risky borrower is more restrictive, this implies that  $(n_S, r_S)$  also satisfies the participation constraint of the safe borrower.

With regards to the incentive compatibility constraint, notice that both contracts lie on  $u_r^1$  which means that a risky borrower would be indifferent between them. Moreover, because of the single-crossing property of indifference curves,  $u_S^1$  the indifference curve of a safe borrower that passes through  $(n_r, r_r)$  stays above  $u_r^1$  for all interest rates less than  $r_r$ , including  $r_S$ . This implies that  $(n_S, r_S)$  is below  $u_S^1$  and is strictly preferred to  $(n_r, r_r)$  by the safe borrower. Therefore, these two contracts satisfy the incentive compatibility constraint as well.

The contract  $(1, \rho/(1 - e^{-r}))$  obviously maximizes the welfare of risky borrowers subject to the constraint of the problem.  $(n_S, r_S)$  is also the utility maximizing contract for safe borrowers, because if it is moved towards the ori-

gin to increase the utility of safe borrowers, the new contract will be preferred to  $(n_r, r_r)$  by risky borrowers, hence violating the incentive compatibility constraint.

This proves that  $(n_r, r_r)$  and  $(n_s, r_s)$  constitute a menu of optimal separating contracts with  $n_s > n_r$  and  $r_s < r_r$ .