Essays on the impact of changing tax rates and welfare spending

Daniel Borbely

Presented in fulfilment of the requirements
for the degree of Doctor of Philosophy

Department of Economics
University of Strathclyde

November 2019

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Acknowledgements

First of all, I am grateful to my supervisors Julia Darby and Graeme Roy for their continuous support during the PhD process. During these last three years they have helped me revise countless error-ridden draft chapters and went out of their way to help me develop both professionally and as a person. This thesis would not have been possible without their enthusiasm and patience.

I am also deeply indebted to David Eiser for being my *de facto* third supervisor and for devoting his time and energy to provide me with useful suggestions and guidance. Most of the research topics that make up this thesis were brought to my attention by David, and he was always there to help when I had trouble navigating through the complexities of Scottish and UK policy matters.

I am also grateful to Markus Gehrsitz for hiring me as a research assistant this year to work on the project *'Class size and human capital accumulation'*. Through our discussions during arduous train journeys between Glasgow and Edinburgh, I have learned a great deal from Markus, and his mentorship during our joint project has made me a better researcher. Along with Markus, I would like to thank my other two co-authors on this project, Gennaro Rossi and Stuart McIntyre, for a productive, intellectually stimulating, and above all, fun experience working together. In particular, special thanks go to Gennaro for putting up with me nearly every day of the last few months. I sincerely hope that this was just the first one of many joint projects.

I would also like to thank my PhD colleagues and friends Ali, Aris, Ayse, Ben, Chirsty, Orion, Oyku, Piero, Ross, and Sharada for all the lengthy lunch breaks and Paesano visits, and for the conversations and shared experiences that made the last three years more enjoyable. I am also grateful to all members of the Strathclyde Economics Department for their help, support, and encouragement.

Finally, special thanks are due to my parents Zsuzsi and Laci, my sister Anna, and my partner Audrey for their continuous love and support.

Synopsis

This thesis examines how changes in tax rates and welfare spending affect the behaviour of markets, households and individuals. In three distinct essays, we make use of applied econometric methods to examine both intended and unintended consequences of changes in tax and welfare policy, their impact on efficiency and equity in the economy, and the way in which they influence people's behaviour. Our findings are intended to contribute to the wider policy debate on the optimal use of tax and spending instruments to achieve policy objectives.

Chapter 1 is titled "*A case study of Germany's aviation tax using the synthetic approach*". This chapter is a case study on the German Aviation Tax, a tax levied on departing passengers from German airports[1]. Aviation taxes receive considerable attention as they have been introduced recently in a number of European countries. Policy makers often motivate these taxes as environmental measures implemented to minimise negative externalities from air travel, or as tools to raise revenues from an otherwise undertaxed sector. Conversely, many experts claim that single-country aviation taxes are inefficient both as environmental measures and revenue-raising tools, due to tax avoidance opportunities presented by tax-exempt airports in neighbouring countries. Critics also point to the adverse effects aviation taxes have on demand, competitiveness and connectivity at domestic air-

[1] This chapter has been published as Borbely, D., 2019. A case study on Germany's aviation tax using the synthetic control approach. Transportation Research Part A: Policy and Practice, 126, pp.377-395.

ports.

Despite their policy relevance, empirical evidence on the effects of aviation taxes is sparse. Only a few studies focus on the issue of aviation taxation, and only a handful of these studies examine the *ex post* effects of newly introduced aviation taxes on passenger flows at tax burdened airports. A key methodological challenge in this small literature stems from the fact that the *ex post* effects of aviation tax changes can only be analysed in quasi-experimental settings, where airports with changes in aviation taxes (treated airports) are compared to a group of airports (controls) not affected by the same policy. Since passenger trends at airports are affected by a wide variety of factors, and are often times highly idiosyncratic, the methodological challenge is to find the appropriate group of control airports to compare treated ones to. Consequently, our main methodological contribution to the literature is to make use of a novel approach, the synthetic control method, to optimise the selection of control airports.

In this chapter, we investigate the impact Germany's aviation tax has had on passenger numbers using the synthetic control method to generate counterfactual passenger numbers for German airports, and for airports outside Germany but near the German border. Our results suggest that passengers engaged in cross-border substitution in response to the aviation tax. Most tax-exempt airports near German borders have made sizeable gains in passenger numbers since Germany introduced its aviation tax. Within Germany there appears to be a clear distinction in the impact on small/regional airports and that on larger hubs - while we find a negative impact on traffic at regional airports, passenger numbers at hubs were largely unaffected (or were even positively affected). From a policy perspective, the finding of a cross-border substitution effect implies that the aviation tax might not be effective in curbing overall emissions from air travel, whilst also leading to lost tax revenues through the displacement of passengers to neighbouring countries. Considering the dual objectives of the policy, the optimal policy outcome from the implementation of AT would be 1) a reduction in overall air travel demand (to cut emissions) and 2) no cross-border substitution effect in order to avoid the loss of the tax base. The findings in this chapter provide evidence to the contrary: German demand, and overall air travel demand, is not reduced to a large extent, but the cross-border effect is sizeable.

Chapter 2 is titled "*Limiting the distortionary impacts of transaction taxes: Scottish stamp duty after the Mirrlees Review*". This chapter investigates the distortionary impacts of transaction taxes through a case study of the Scottish residential property market. Tax on property transactions, also known as stamp duty, is a common form of property taxation but not a very popular one. Economists consider these taxes inefficient for a number of reasons. First, they reduce the incentive to engage in mutually beneficial property transactions, leading to welfare loss as properties are not owned by the individuals who assign the most value to them. Second, they increase the tax burden on those who, for whatever reason, trade housing more often. Lastly, by increasing the cost of mobility, transaction taxes make it more difficult for homeowners to relocate in search of appropriate employment.

In line with these efficiency concerns, recent policy changes in Scotland were aimed at reducing distortions from the transaction tax system by 1) simplifying the structure of the tax system and 2) reducing the tax burden for the majority of home buyers (and increasing it for those aiming to buy expensive properties). Although recent studies provide credible empirical evidence on the distortions created by transaction tax regimes, the key contribution of our study is that it is the first one to study policy changes aimed at reducing said distortions. From a policy perspective this is important since governments might struggle to replace transaction tax regimes for budgeting or administrative reasons, leaving sensible reform as the only feasible policy option to mitigate efficiency concerns.

To study the property market impacts of transaction taxes and assess the effects of recent reform, this chapter makes use of four sources of variation in transaction tax rates present in recent Scottish tax systems: price notches, kink points, time notches and a shift to a more progressive transaction tax regime. Our analysis makes use of bunching estimator and difference-in-differences approaches to examine how variations in transaction tax rates influence transaction activity on the Scottish property market. Our results show that both kinds of notches have a distortionary impact that is sub-optimal. The Scottish Government's recent reforms had a positive impact through replacement of price notches with kink points, but time notches emerged allowing other distortions to persist. Through a comparison of price ranges affected and unaffected by the policy changes, we also find that progressive reform had a significant (and long-term) positive impact on transaction activity in the market segment where tax rates were reduced. On the other hand, the higher end of the market, where tax rates increased, was mostly unaffected by progressive reform with the exception of the market for very expensive properties, where a negative effect is identified. Implications of our findings are that if governments want to make transaction tax regimes more efficient, progressive taxation might be a good way to limit their distortionary impact, whilst also encouraging transaction activity in the lower end of the market.

Chapter 3 is titled "*The impact of housing subsidy cuts on the labour market outcomes of claimants: Evidence from England*". This final chapter looks at the labour market impacts of reduced housing subsidies. Housing subsidies are aimed at helping low-income individuals afford appropriate housing, but are costly to offer. Also, in the view of some experts and policy makers, housing subsidies reduce work incentives and/or work effort among claimants. Economic theory is ambiguous on the likely labour market impacts of changes in housing subsidy entitlements. To date, there is no credible empirical evidence linking housing subsidy cuts to positive changes in labour market outcomes. Nonetheless, policy makers often motivate cuts to subsidies as a way to encourage claimants to seek employment (or put in more work effort).

This chapter investigates the labour market impacts of recent housing subsidy cuts in England that were justified using a similar argument to that discussed above. Our investigation makes use of individual-level survey data from the Understanding Society survey. Using this data set allows us to track subsidy claimants over time and control for a wide range of individual and household characteristics. Our identification strategy relies on the fact that, within the time period investigated, the subsidy cuts were only implemented for claimants renting from one segment of England's rental sector (the private rental sector) while claimants renting from other segments (the social rental sector) were unaffected. We utilise this variation in exposure to the subsidy cuts within a difference-in-differences framework to estimate the effect of reduced housing subsidies on the labour market outcomes of claimants. We find no evidence of a change in labour market outcomes for those affected by subsidy cuts, indicating that, at least on aggregate, the reforms did not succeed in encouraging employment among claimants. These null findings should be of considerable interest to policy makers who often assume a causal link between the generosity of benefit entitlements and the work effort/incentives of claimants. More specifically, our findings suggest that as a policy instrument, cuts to housing subsidies may not be effective in generating efficiency gains through increased labour market participation or work effort.

Contents

List of Figures

List of Tables

Chapter 1

A CASE STUDY OF GERMANY’S AVIATION TAX USING THE SYNTHETIC CONTROL APPROACH

1.1 Introduction

Aviation tax (AT) regimes in Europe receive considerable attention. Germany's AT was introduced on 1^{st} January 2011 and remains payable on departures from all German airports at a cost of €7.50, €22.43, or €42.18 depending on the distance flown[2]. While often motivated as environmental taxes, the common perception is that revenue raising is the key driver of ATs (Zuidberg, 2015). In spite of this, experts often question the efficacy of single-country ATs as environmental measures or revenue-raising tools, in large part due to the cross-border tax avoidance opportunities associated with them (Zuidberg, 2015; Krenek and Schratzenstaller, 2017). Airports and airlines often spearhead campaigns for abolition of ATs, citing adverse effects on demand, competitiveness and connectivity[3]. Germany's Economy Minister is reportedly swaying toward abolition for these reasons[4].

In this chapter, we investigate the impact that Germany's AT has had on passenger numbers using German airports and airports outside Germany but near the border (henceforth referred to as bordering airports). Specifically, we use the synthetic control method of Abadie et al. (2010) to construct counterfactual series for each airport of interest, the counterfactual series representing passenger numbers under the alternative scenario that AT was never introduced. The impact of AT is then estimated as the gap between actual passenger numbers since 2011 and the counterfactual numbers.

We estimate changes in passenger numbers that can be attributed to AT for German airports and for bordering airports outside of Germany. Airports are modelled separately to see whether the effects of the AT might differ across airport types. Results indicate more passengers used bordering airports after the introduction of German AT, while most German airports, with the exception of hubs, saw a negative impact of AT on passenger

[2] See Deutsche Bundesregierung (2012).
[3] See Edinburgh Airport (2015).
[4] See DW (2017).

numbers.

Our main contribution is to the literature on the impacts of AT. Case studies focusing on the Dutch AT (introduced in 2008 and abolished a year later) find that the aviation tax led to a significant reduction in passenger numbers at Dutch airports, an increase in passengers at airports in bordering countries (Gordijn and Kolkman, 2011), and a reduction in tourism (Mayor and Tol, 2010)[5]. For the United Kingdom, Seetaram et al. (2014) find that the UK aviation tax (Air Passenger Duty, or APD) was ineffective in reducing demand for air travel and passengers were prepared to pay higher prices for flight tickets. Similarly, Mayor and Tol (2007) find that doubling the UK aviation tax did not lead to a reduction in demand, only a substitution from short-haul to long-haul flights. For Australia, Forsyth et al. (2014) estimates that the introduction of the aviation tax will lead to sizable net losses to the tourism industry. Finally, contemporaneously to our study, Falk and Hagsten (2019) assess the impacts of the recently introduced AT in Germany and Austria on passenger numbers at both domestic and bordering airports. They use a panel data difference-in-differences (diff-in-diff) approach with untaxed European airports as the comparison (control) group in order to estimate the effects of AT on passenger numbers. They also distinguish between hubs, low-cost and bordering airports: their main finding is that the AT led to an overall reduction in passenger numbers at Austrian and German airports, but that these results are mostly driven by airports that predominantly serve low-cost airlines. They find no significant impact from the AT on passenger numbers at bordering airports.

Case studies that have assessed the impacts of ATs on passenger numbers have either used a linear extrapolation method (Gordijn and Kolkman, 2011) or panel diff-in-diff (Falk and Hagsten, 2019) to estimate counterfactual passenger numbers at airports, where

[5] While the analysis of the Dutch case by Gordijn and Kolkman (2011) finds similar effects from AT to the ones uncovered in this paper, their analysis is largely descriptive and causal inference of the effects is not established.

the counterfactual approximates passenger numbers under a scenario of no aviation tax. Through our empirical strategy, we contribute to the literature by employing a novel approach to estimate counterfactual passenger numbers at airports. Our empirical approach uses the synthetic control method (Abadie et al., 2010) to estimate counterfactual passenger numbers, and can provide robust estimates of counterfactuals for a number of reasons. First, this method optimises the selection of comparison (control) airports so that counterfactual passenger numbers are based on those control airports most similar to the treated airports in terms of passenger trends. Second, contrary to panel diff-in-diff estimations, synthetic control estimates do not rely on the assumption of parallel trends for treated and control groups (Billmeier and Nannicini, 2013). This approach is particularly suitable for our analysis as it allows us to construct reliable counterfactuals using aggregate level data. Indeed, according to Abadie et al. (2015) the most relevant application of the synthetic control method is for comparative case studies that use aggregate data.

Through our findings we also contribute to the literature on the determinants of air passenger demand and airport choice (see for example Graham, 2000; Valdes, 2015; or Jankiewicz and Huderek-Glapska, 2016). Only a few studies (Pels et al., 1998; Steverink and van Daalen, 2011) consider the impact that aviation taxes have on the airport choices of passengers. These studies are however mostly theoretical in their approach. This chapter therefore contributes to the literature by providing empirical evidence that passengers are highly responsive to the introduction of AT, and by finding some evidence (albeit only at the aggregate level) that passengers change airport choices in response to these taxes.

Furthermore, our analysis contributes to the literature on cross-border shopping (Joossens and Raw, 1995; Nielsen, 2001; Asplund et al., 2007). This strand of the tax policy literature focuses on the way tax differences across borders affect the preferences and choices of consumers. Our results contribute to the literature by finding evidence for a specific case of cross-border shopping: our results show that AT differences across borders can have a noticeable impact on air passenger movements between the bordering countries. More specifically, we find that the introduction of the AT in Germany can be associated with a substantial increase in passengers at tax-free airports near the German border.

The remainder of the chapter is organised as follows. Section 1.2 summarises the relevant conceptual and policy background. Section 1.3 describes the data and outlines the empirical approach. Section 1.4 shows the results and assesses the plausibility of our findings. Section 1.5 provides a discussion of our findings. Section 1.6 concludes.

1.2 Conceptual and Policy Background

The German aviation tax (AT) was introduced on 1^{st} January, 2011. The AT is charged on passengers departing from German airports. There are three different tax bands charged based on distance flown with rates set at €7.50, €22.43, and €42.18 for short, medium, and long-distance flights, respectively[6]. According to forecasts by Berster et al. (2010), the number of passengers expected to belong to each tax category was (at the time of introduction) 62.3, 2.9, and 8.9 million, respectively.

We estimate that roughly 84% of passengers can be expected to pay a unit tax of €7.50, 4% of passengers can be expected to pay €22.43, and 12% of passengers can be expected to pay €42.18[7]. Consequently, passengers on short and long-distance flights contribute the vast majority of tax revenues from AT. The average passenger departing from a German airport faced a tax increase of €12.26[8].

[6] See FCC Aviation (2011) and Deutsche Bundesregierung (2012).

[7] Unfortunately no data are available to estimate these numbers for all of our sample years. The estimation above is based on 2008 data from Berster et al. (2010).

[8] This is the weighted average of the unit taxes considering the shares of passengers traveling under each distance band.

In Europe, only a few countries charge AT directly on air passengers, although nearly half of European countries impose either direct or indirect taxes on aviation (European Commission, 2019). When comparing effective tax rates under these tax regimes, German AT rates stand out as one of the highest in Europe, and tax rates are especially high compared to most bordering countries[9].

The introduction of German AT was justified as an environmental measure, however, experts believe that the German government's main objective with the tax was to consolidate its budget (see Krenek and Schratzenstaller, 2017)[10]. The government's initial expectation was to raise €1 billion annually from AT (Berster et al., 2010; Steppler, 2011). Upon the introduction of the tax, some experts warned (see Steppler, 2011) that the German AT was likely to have the same adverse effect on passenger numbers that had been experienced following the introduction of the Dutch version of the tax a few years earlier. The AT remains controversial to this day with frequent calls for its abolition by industry participants and policy makers.

Aviation taxes may impact both the supply (airlines' supply of flights, destinations, and frequencies) and the demand (passengers' demand for airline services) side of the aviation industry. On the supply side, upon the implementation of AT, airlines initially decide to what extent they should pass the cost of these taxes to passengers: their choice is whether to keep ticket prices constant (and absorb the tax themselves) or pass on the tax in the form of higher ticket prices. This decision largely depends on market conditions and the type of strategy airlines employ (Zuidberg, 2015): profit-maximising airlines are more likely to pass taxes on in comparison with growth-maximizing airlines (these will likely absorb the tax to maintain traffic).

[9] European comparisons rely on exchange rate adjusted average effective AT rates in each country. The highest AT rates in Europe are charged in the UK (European Commission, 2019)

[10] In the German Aviation Tax Act 2011, it is stated that AT should "provide an incentive for behaviour which is more appropriate to the needs of the environment".

Airlines might also adjust their networks in response to the introduction of aviation taxes. They might choose to serve different airports and provide lower flight frequencies at certain airports (Zuidberg, 2015). These adjustments are more likely to be made by low-cost airlines as these airlines are more flexible, and can move airports at relative ease in comparison to traditional carrier airlines as their operations are not tied to large hub airports (Thelle and la Cour Sonne, 2018). For this reason, in the case of the German AT, *ceteris paribus*, we should expect larger responses to aviation taxes at smaller regional airports that predominantly serve low-cost airlines than at large hubs. According to Zuidberg (2015), after the implementation of the German AT, the low-cost carrier Ryanair announced that it would cut capacity at a number of German regional airports (for example Bremen or Frankfurt Hahn) it used for its operations. Supply side responses of this kind will likely also impact on passenger demand at these airports.

The network adjustment responses of airlines will be partly in anticipation of (or in response to) changes in passenger demand in response to the tax changes (assuming at least some level of pass-through). Changes in passenger demand in response to AT will largely depend on the price elasticity of demand. Price elasticities for air travel might be highly heterogeneous across consumer groups, airports and types of travel (see for example Brons et al., 2002 or Hofer et al., 2008). Specifically, short-haul and leisure passengers are likely to be more sensitive to price changes than long-haul and business passengers (Brons et al., 2002; Hess and Polak, 2005; Morlotti et al., 2017). Short-haul passengers might be more price sensitive due to the ease with which they can substitute to other forms of transportation (cars, railway). Once again, short-haul and leisure passengers are likely to provide a large percentage of the traffic at smaller regional airports, while there should be more long-haul and business passengers at hubs, so we should expect a larger demand side response to AT at small regional airports and a lesser response at hubs.

Another possible substitution response by passengers is to start using nearby airports in bordering countries where no aviation tax is in place. In the case of the Dutch AT there is some evidence that the aviation tax prompted Dutch passengers to substitute to nearby German and Belgian airports (Gordijn and Kolkman, 2011). It is also likely that airlines respond to this substitution effect by adjusting their network and relocating some services to AT exempt bordering airports. In the German AT case, we would expect that the presence of viable AT exempt airports near the German border (for example Eindhoven, Charleroi, Luxembourg or Basel) would induce a substitution effect (even though Falk and Hagsten (2019) found no significant evidence of this in their analysis). The leakage of air passengers induced by the substitution to bordering airports might also be problematic for governments from a policy perspective: it leads to lower tax revenues but does not significantly reduce the overall number of air passengers, so no improvement in environmental outcomes can be claimed.

1.3 Data and Identification Strategy

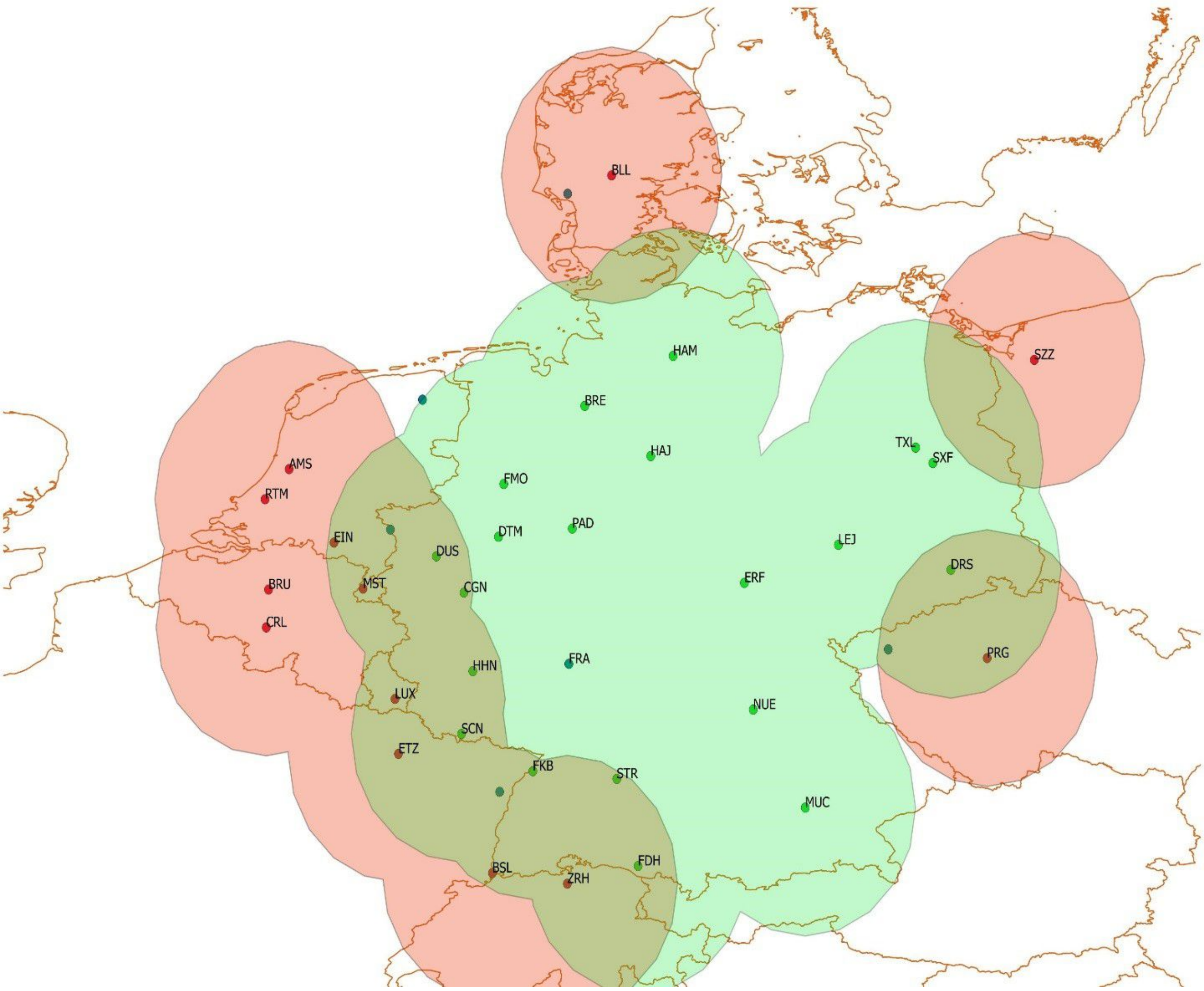
We assess the impact of the German AT on passenger numbers by estimating counterfactual series of passenger numbers for each treated airport – the counterfactual numbers correspond to a scenario where no AT was introduced. The choice of 'treated' airports, for which we believe AT may have had an impact, are German international airports and bordering airports (located outside Germany but within two hours driving time). Our calculations for driving time are based on Google Maps data. Our choice of catchment area is based on the likely cost considerations of passengers: per passenger fuel costs for a two-hour car journey on a motorway near German border regions are similar in value (according to Journey Price Calculator[11]) to the average tax savings from avoiding AT. The choice of catchment area only affects our selection of possible 'treated' bordering airports

[11] Journey Price Calculator is a free Internet tool to estimate the cost of car journeys. See Journey Price Calculator.

and has no impact on estimations for other airports[12]. Including bordering airports in the analysis allows us to investigate potential spill-over effects from AT. Treated airports are shown in Figure 1.1 along with an indication of catchment area, represented by a circle with a radius of 150km (corresponding to approximately two hours of driving time). Blue dots represent airports excluded due to data/model limitations. Our sample includes 21 treated airports in Germany, and 13 treated airports near the German border (but outside of Germany). To assess why AT might have had a different impact at some airports in comparison to others, we estimate separate models for each airport. For treated airports, we compare actual (observed) passenger numbers to counterfactual ones after the introduction of AT to assess the impact of the tax.

[12] There are no potential 'treated' bordering airports near the catchment area cutoff and therefore the selection of treated airports is not sensitive to marginal changes in the radius of the catchment area. It is also worth noting that some airports may be more easily reached through Germany's highly developed train network. Train journeys may offer time savings when compared to car journeys on certain routes but should not alter airport catchment areas to a large enough extent to allow us to include more bordering airports.

Figure 1.1: *Map of Airports and Catchment Areas*



Notes: German airports and catchment areas are shown in green; bordering airports/catchment areas in red; and points with no label indicate airports excluded from estimation. Information on airport codes can be found in Tables 2-3 in the Appendix A.

1.3.1 Synthetic control method

To estimate counterfactual passenger numbers, we use the synthetic control method (Abadie et al., 2010). The synthetic control method constructs counterfactuals (also referred to as synthetic controls) using a weighted average of 'control' airports. Control airports are airports where no changes in AT took place during our sample period[13]. Formally, the synthetic control method as applied to the German AT case can be described as follows.

Let J be the number of airports in our sample, where $j = 1$ is a specific treated airport. In order to allow for a meaningful comparison, the sample J contains airports similar in characteristics to the treated airport (see Section 1.3.2 below for more detail).

The airports ($j = 2, \ldots, J$) are included in the group of possible control airports. The counterfactual (synthetic control) is a weighted average of control airports and can be represented as a ($Jx1$) vector of weights $W = (w_2, \ldots, w_{J-1})$, where $w_2 + \cdots + w_{J-1} = 1$ and $0 \leq w_j < 1$.

Furthermore, let X_T be a $(kx1)$ vector containing the values of pre-AT characteristics of the treated airport, and let X_C be a (kxJ) matrix containing values for the same characteristics for control airports. These characteristics include predictors of passenger numbers (covariates such as regional purchasing power or flight ticket price inflation) and also pre-AT passenger numbers at airports. We would like to match treated and control airports based on the values of predictors so that our counterfactual best resembles the treated airport in terms of pre-AT characteristics. The difference in pre-AT characteristics between the treated airport and the counterfactual is equal to $X_T - X_C W$. The optimal counterfactual W^* (a weighted average of control airports) selected through the synthetic control procedure is the one that minimises the size of this difference.

[13] Information on aviation tax regimes is from EBAA (2015).

Applying W^* to post-AT passenger numbers then gives the counterfactual passenger number $P_C W^*$, which is equal to passenger numbers at selected control airports multiplied by their optimal weights. The effect of AT on passenger numbers at the treated airport is then given by $P_T - P_C W^*$, that is, the post-AT difference between actual passenger numbers at the treated airport (P_T) and counterfactual estimates ($P_C W^*$).

In summary, the synthetic control method chooses control weights so that pre-AT passenger numbers and covariate (predictors) values for the counterfactual are similar (on average) to those at the treated airport. The covariates also control for variation arising from other factors that affect passenger numbers (regional and macroeconomic variables). The synthetic control approach therefore optimises the selection of control airports: to construct counterfactual passenger numbers the procedure finds the convex combination of control airports that provide the 'best fit', that is, the fit between passenger numbers at treated airports and their synthetic control, pre-AT.

As mentioned above, in this chapter we construct counterfactuals using separate synthetic control estimations for each treated airport[14]. For these airports, we then compare actual passenger numbers to counterfactual ones after the introduction of AT to assess the impact of the tax.

According to Abadie et al. (2015) synthetic control should be used only when there is a sufficiently large number of pre-treatment periods for the optimisation of the control unit selection. While the number of pre-treatment periods we have in our analysis (eight) is certainly not large, there is no precise recommendation for the number of pre-treatment periods, and several recent examples of research using synthetic control used fewer pre-treatment periods (Munasib and Rickman, 2015; Zhang et al., 2016; or Chamon et al.,

[14] Synthetic control estimations in this chapter are implemented using the 'synth' package in Stata.

2017)[15].

The synthetic control approach is similar to the panel diff-in-diff methodology often used for policy evaluations and also employed in Falk and Hagsten (2019) for their analysis of the impact of Austrian and German AT. We believe however that the synthetic control approach has some notable advantages over panel-diff-in-diff in the context of the present analysis.

First, synthetic control relies on considerably weaker identifying assumptions in comparison to panel diff-in-diff methods (Billmeier and Nannicini, 2013). In particular, panel diff-in-diff estimations rely on the assumption of parallel trends in outcomes for treated and control units. In our example, under this assumption, passenger numbers at treated (taxed) airports would have to follow the same pre-AT trends as control (untaxed) airports - post-AT differences in passenger numbers could then be contributed to the presence of AT. For the German AT case, this assumption is difficult to validate due to the high degree of heterogeneity in passenger trends across different airports: passenger numbers are determined by a number of factors and conditions (see Hess and Polak, 2005) and it might not be safe to assume that all differences (even conditional on covariates) between treated and control airports are a result of tax changes. On the other hand, synthetic control does not rely on the assumption of parallel trends as it optimises control airport selection on the basis of pre-AT fit in passenger trends between treated and control airports.

Another advantage of synthetic control in comparison to panel diff-in-diff is that it can account for the presence of time-varying unobservable confounders (Billmeier and Nannicini, 2013). This provides a good way to minimise the possibility of omitted variable bias and control for the fact that some shocks (captured by our covariates) to passenger numbers might not occur until after the introduction of AT.

[15] We are unable to use earlier data on our sample airports as passenger data are not consistently available before 2003. For example, Eurostat only has data on passenger numbers at German airports from 2003 onwards.

There are limitations to the synthetic control approach that must be considered when applying this method. First synthetic control estimates may be biased by idiosyncratic shocks to the outcome variable that control units and covariates are unable to capture (Abadie et al., 2015). In the case of German AT, estimates may be biased due to idiosyncratic country, or airport, specific shocks to passenger numbers at German airports. Examples of such shocks to passenger numbers could include terrorist attacks, natural disasters and/or major sporting events. In the context of this chapter, a notable concern is that the impact of the 2006 FIFA World Cup, and particularly its final in Berlin, on passenger numbers might interfere with our counterfactual estimations. To ensure that we do not draw conclusions based on biased estimates, we check the validity of our counterfactuals through several sensitivity checks which are presented in Section 1.4 below. Another limitation of the synthetic control approach is that it does not allow for large-sample asymptotic inference of estimates (Billmeier and Nannicini, 2013). Nonetheless, inferences can be made using placebo and robustness tests (Abadie et al., 2010) which are reported in Section 1.4 of this chapter.

1.3.2 Control airport selection

As mentioned in Section 1.3.1, we estimate separate synthetic control models for each treated airport. This is done in order to 1) optimise counterfactual estimations for treated airports through the selection of appropriate controls and 2) to assess the heterogeneity of AT impacts at different airports. In order to avoid overfitting (see Abadie et al., 2015), a different set of possible control airports is selected for each treated airport. Overfitting can occur when the control group contains airports that are dissimilar to the treated airport in their main characteristics and the counterfactual estimation is optimised based on artificially matched airports as a result of idiosyncratic variation in passenger numbers.

Criteria employed in selecting control airports include: located in a country with no change in AT over the period; but otherwise similar characteristics to the specific treated airport. Similarity is considered in terms of passenger numbers at treated and control airports. An airport is not included in the possible control group of a treated airport if maximum passenger numbers (over the pre-AT period) at the treated airport were at least twice as large (or half as large) as those at the given control airport. Full details on treated airports and their controls are provided in the Appendix A, along with the selected weighted averages used in order to create counterfactual passenger number series.

Following Billmeier and Nannicini (2013), two sets of estimates are constructed for each treated airport: set A use control airports from countries surrounding Germany under the assumption that treated and control units are likely to face similar macroeconomic and regional shocks; while set B use a larger number of control airports from across the European Economic Area (EEA). The preferred results are those from whichever set of estimates provides the best fit for the pre-tax period (as in Ormaechea et al., 2017).

1.3.3 Covariate selection

In our analysis, we include a number of covariates to control for factors affecting passenger trends at airports. The covariates included are: purchasing power per capita in euros (in constant prices) at the NUTS2 regional level, to control for the impact of macroeconomic shocks; lagged passenger numbers, to capture airport specific trends; and flight ticket price inflation in the country of the airport, to control for changes in ticket prices on the aggregate level. First-differences are also used to control for airport fixed effects. Annual data on passenger numbers using each airport over the period 2003-2015 are from

Eurostat[16]. Data on purchasing power per capita, and flight ticket price inflation are also from Eurostat.

1.3.4 Data restrictions

To avoid biased estimates, we need to make restrictions to our data. Estimation bias could arise from including treated airports in the group of possible controls: under such circumstances counterfactual passenger numbers might also be affected by AT changes. Austrian airports are therefore excluded from the group of possible control airports, since Austria also introduced an AT in 2011. Furthermore, contrary to Falk and Hagsten (2019), we do not 'treat' Austrian airports and investigate their AT's impact on passenger numbers. This is because the Austrian AT charged different rates from the German version of the tax – our preference is to keep the treatment homogenous so that we can assess the heterogeneity of treatment effects across airports and relate these differences to airport characteristics.

In addition, the Netherlands introduced an AT in July 2008 but abolished it a year later. It is possible that some of the longer-term effects (see Zuidberg, 2015) of the abolition of Dutch AT will coincide with the introduction of the German aviation tax – this is a limitation of our analysis[17]. Nonetheless, we do not exclude Dutch airports from the group of treated airports as they provide some of the best substitution opportunities for German passengers after the introduction of AT – if we observe substitution effects after AT these are likely to concern Dutch airports close to the German border. Furthermore, German AT was introduced more than a year after the abolition of Dutch AT, our estimations there-

[16] These data concern both departing, arriving and transfer passengers, which is a limitation since the tax is charged only on departing passengers.

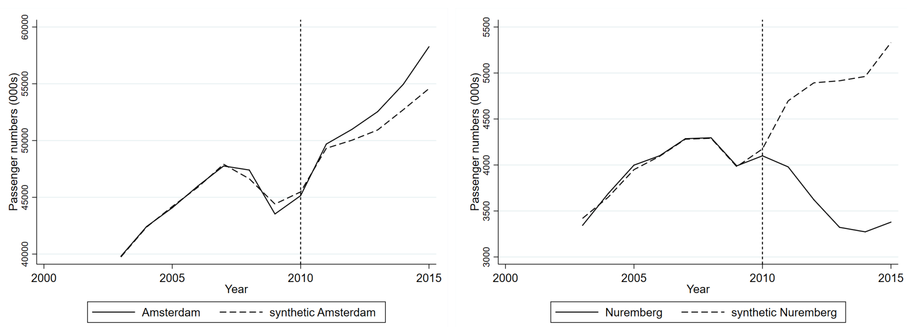
[17] This is a limitation for the estimations that concern Dutch airports. As Dutch airports are 'treated' and are therefore not in the control groups of other airports, estimations for other airports are unaffected by our decision to include them in the sample.

fore likely capture most of the impact of the abolition on passenger numbers during this period.

1.4 Results

The full synthetic control results are provided in the Appendix A. Here we show results for two airports by way of examples. Figure 1.2a and Figure 1.2b plot actual and counterfactual (synthetic) passenger numbers for Amsterdam and Nuremberg respectively. The impact associated with the introduction of AT corresponds to the vertical difference (gap) between the actual and counterfactual time trends after 2010. The imposition of AT in 2011 is associated with increased passenger numbers relative to the counterfactual in the case of Amsterdam, and decreased numbers in the case of Nuremberg.

Figure 1.2: *Examples of synthetic control estimates*



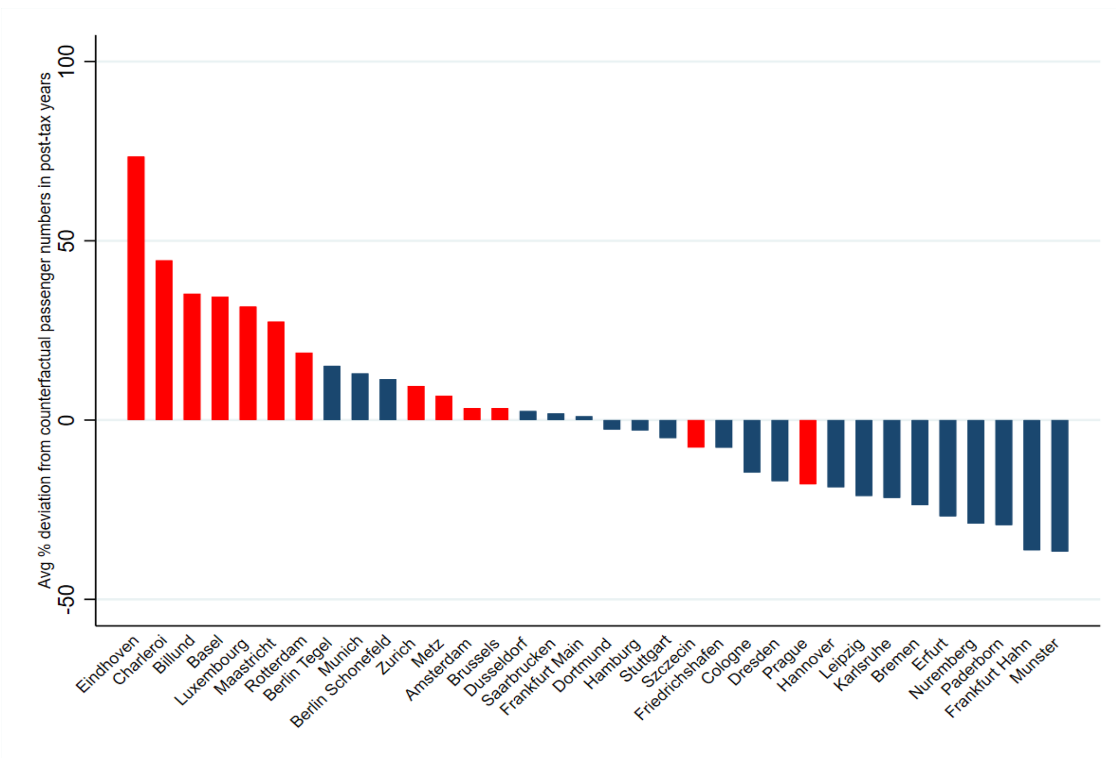
(a) Amsterdam (b) Nuremberg

Figure 1.3 summarises the full set of results for the treated airports, by plotting the post-tax percentage deviations of actual passenger numbers as compared to the counterfactual. The plotted figures are each airport's average percentage change in passenger numbers over the post-tax sample, i.e. 2011-2015, in each case.

Recall from Section 1.2 that different airport types (specifically hubs as distinct from regional airports) might respond differently to the implementation of AT. For this reason, Figure 1.4 provides a summary of the results grouped by (treated) airport type. Airports are designated as hubs based on their size and the airlines that use them as the centre for their operations and networks: the four largest German airports (Frankfurt Main, Munich, Berlin-Tegel and Dusseldorf) are obvious choices as they handle a significantly larger number of passengers in comparison to other airports and also serve as the hubs of a number of airlines. We include Hamburg Airport in the hub category as it serves a considerably larger number of passengers in comparison to the next busiest German airport (Cologne). All non-hub airports in Germany are in the regional airports group, while bordering airports are in a separate group regardless of their size. Finally, we create a separate 'low-cost' airport group to assess whether airports serving predominantly low-cost airlines were impacted differently by the AT introduction. Low-cost airports are those airports where at least 50% of the services are provided by low-cost airlines[18]. This category contains almost all German non-hub airports, with the exception of Stuttgart.

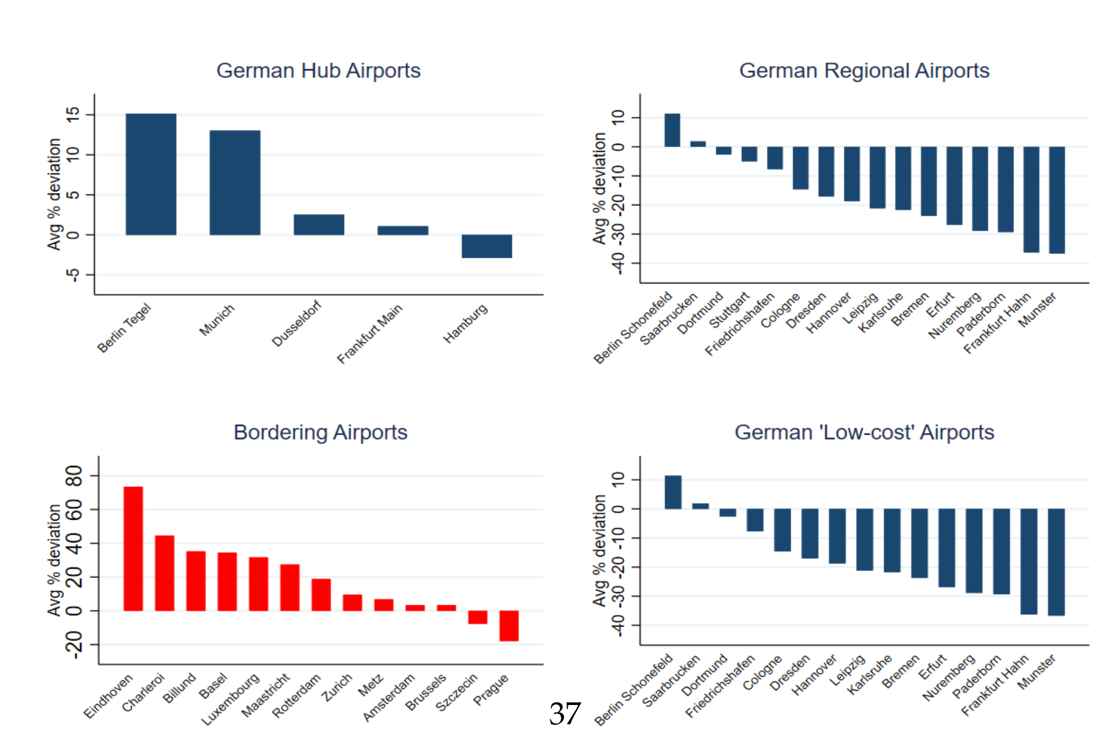
[18] Information on services is collected from airport websites. This information is not available for previous years so 2018 data are used.

Figure 1.3: *Average Post-Tax Percentage Deviations from Counterfactual Passenger Numbers (Treated Airports)*



Notes: The chart shows the average of the yearly deviations of actual passenger numbers from counterfactual ones. Averages are taken over the post-tax years for each airport. Bordering airports are marked red.

Figure 1.4: *Average Post-Tax Percentage Deviations from Counterfactual Passenger Numbers - by Airport Type*



Notes: The charts show the average of the yearly deviations of actual passenger numbers from counterfactual ones. Averages are taken over the post-tax years for each airport.

Most German airports appear on the right of the Figure 1.3, indicating estimated losses of passengers since AT's introduction. Most bordering airports, shown in red, made gains. A small number of German airports, notably Berlin Tegel, Dusseldorf and Munich saw gains in passenger numbers relative to the no tax counterfactual. In Figure 1.4, it is also clear that most regional airports in Germany (most of these airports serving low-cost airlines) lost passengers (in comparison the counterfactual scenario) since the introduction of AT.

Aggregating the results shown in Figure 1.3 for German airports provides us with an estimate of 4 million passengers lost (annually) compared to the counterfactual scenario under the assumption of no AT. This is roughly 2% of annual passenger numbers at these airports so indicates rather a small aggregate (net) effect on passenger demand at German airports.

1.4.1 On the plausibility of the findings

To check the plausibility of our findings we: provide information on pre-AT fit of models; explore the sensitivity of results to the choice of control airports; and provide information on the significance of our estimates based on placebo tests.

The fit of synthetic control models in the pre-AT period is measured using Root Mean Squared Prediction Errors (RMSPE) - constructed from the squared difference between actual and counterfactual passenger numbers for each pre-AT year, averaged across the available pre-tax years. The normalised RMSPE adjusts for airport size so is expressed as a % of that airport's passenger numbers in 2010; a figure in excess of 5% is indicative of poor fit and signals that estimates of the impact of AT must then be treated with caution.

The check on the sensitivity of results considers whether estimated impacts of AT are affected by the inclusion/exclusion of particular airports in the set of controls. Recall from Section 1.3.2, that we estimated separate synthetic control models for each treated airport using different sets of control airports (set A and set B)[19]. Similar estimated impacts from the two sets are indicative of robustness (see Table 1.1, column 2).

Asymptotic inference is not possible using the synthetic control method. Inference on significance of estimated effects therefore comes from placebo tests, following Abadie et al. (2010). First the AT treatment is assigned to each control airport and impacts estimated. Since no AT was actually introduced for these airports we expect any estimated impacts to be small and random. Essentially, we have confidence in our results for treated airports if their estimated post-tax gaps are large relative to those generated in the placebo tests[20].

To evaluate this, we use RMSPE ratios to construct p-values. RMSPE ratios are measured as the post-tax gap (between actual and counterfactual passenger numbers) divided by the pre-tax gap. These ratios indicate the extent to which post-tax gaps are large in comparison to the pre-tax fit of our counterfactuals. Each p-value then indicates the likelihood that a randomly selected RMSPE ratio from the sample of placebo tests is larger than that of the given treated airport. It is simply the number of RMSPE ratios from the placebo group that exceed the ratio for the treated airport, and divided by the number of control airports in the group. For example, if the treated airport's RMSPE ratio is larger than the ratio for all of its control airports (say there are 50 of them), the p-value is going to be equal to 0/50 = 0. Significance can be interpreted as indicating that AT had an unusually large effect on a treated airport relative to the placebo effects for control airports. Recall though that inference must be predicated on well-fitting models that are reasonably robust to the choice of controls. Table 1.1 summarises our findings.

[19] The synthetic control results presented in the Appendix A show estimates from the set of controls that provides the better pre-AT fit. Figures showing results for estimations using the other set of control variables can be found in the Online Appendix (https://sites.google.com/view/danielborbely/research).

[20] Graphical representations of the placebo tests for each treated airport can be found in the Online Appendix.

Most of our airport estimates are based on synthetic control models that achieve a good fit prior to the introduction of AT and are robust. Some of these results are not significant however. This is likely a consequence of relatively small effects associated with AT at these airports (for example Frankfurt Main), or large idiosyncratic effects coinciding with AT at some of the control airports. As the results for these treated airports are robust and are based on reliable counterfactual estimations (good pre-AT fit), we should not reject them despite the lack of significance.

In Table 1.1, airport results with pre-estimation errors of over 5% are based on ill-fitting synthetic control models. Counterfactual estimations for these airports are unreliable as counterfactual passenger numbers already substantially deviate from actual numbers before the introduction of AT[21] The small number of airports falling under this category tend to be small regional airports, with low annual passenger numbers. It is possible that passenger number changes at these airports are too idiosyncratic to be modelled appropriately. Since these airports serve very few airlines and destinations, a single airline changing routes or schedules might have a substantial impact on passenger numbers.

[21] The estimates for these airports are still included in the results summarised in the last two sections. Nonetheless, the estimates for these airports are imprecise and related findings do not hold up to statistical scrutiny. These estimates do not affect the estimates and findings for other airports however, and do not change the general conclusions we can draw from our findings.

Table 1.1: *Summary of robustness/inference measures*

Airports	**Pre-estimation error (% of passengers)**	**Sample Robustness Check**	**RMSPE Ratio**	**p-value**
German hubs				
Berlin Tegel	3.60	✓	4.61	0.273
Dusseldorf	1.49	✓	2.92	0.700
Frankfurt Main	0.53	✓	2.17	0.818
Hamburg	1.62	✓	4.16	0.706
Munich	0.53	✓	24.56	0.000***
German regional airports				
Berlin Schonefeld	5.89	✓	1.75	0.682
Bremen	2.44	✓	12.91	0.032**
Cologne	0.61	✓	27.17	0.000***
Dortmund	3.65	✓	1.45	0.914
Dresden	0.55	✓	37.49	0.000***
Erfurt	16.68	✓	1.58	0.726
Frankfurt Hahn	4.11	✓	11.23	0.076*
Friedrichshafen	2.07	✓	3.71	0.313
Hannover	4.00	✓	6.17	0.238
Karlsruhe	5.83	✓	4.55	0.600
Leipzig	2.62	✓	9.07	0.152
Munster	1.33	✓	33.15	0.000***
Nuremberg	0.81	✓	43.62	0.000***
Paderborn	6.02	✓	5.95	0.393
Saarbrucken	7.88	✓	1.26	0.813
Stuttgart	1.37	✓	4.19	0.474
Bordering airports				
Amsterdam	0.68	✓	22.13	0.000***
Basel	3.93	✓	10.36	0.038**
Billund	3.41	✓	8.32	0.111
Brussels	1.53	✓	4.26	0.818
Charleroi	6.59	✓	4.49	0.300
Eindhoven	4.77	✓	169.09	0.000***
Luxembourg	2.80	✓	11.71	0.057*
Maastricht	18.32	✓	1.97	0.643
Metz	5.32	✓	1.48	0.867
Prague	1.99	✓	10.76	0.143
Rotterdam	8.71	✓	2.91	0.733
Szczecin	10.71	✓	1.52	0.714
Zurich	3.02	✓	3.06	0.636

Notes: ***Significant at 1% level, **5%, *10%. The synthetic control model for a given airport is considered ill-fitting when pre-estimation error (RMSPE) is higher than 5% of passenger numbers in 2010. The sample robustness check relies on the comparison of estimates that use different samples of control airports.

1.5 Discussion

The findings set out in Section 1.4 are consistent with the likely behavioural responses of agents to increases in AT.

a) That bordering airports are estimated to have benefited from AT is consistent with passengers substituting to alternative airports to avoid ticket prices that incorporate AT. Such effects will be strongest when German and non-German airport catchment areas overlap (for example in the case of Eindhoven – EIN or Basel - BSL, see Figure 1.1). Our findings of sizable substitution effects at most bordering airports contrast with Falk and Hagsten (2019), who find no significant impact on bordering airports from German and Austrian AT using a panel diff-in-diff approach. It is possible that differences in results arise from methodological differences: while our approach allows us to observe airport-specific treatment effects at each bordering airport, in Falk and Hagsten (2019) bordering airports are pooled together in the panel diff-in-diff analysis potentially allowing outlier airports (where AT had no effect or a negative effect) to put a downward bias on the overall treatment effect. Outliers are also present in our sample: in the case of two bordering airports (Szczecin and Prague) decreased passenger numbers can be associated with the introduction of AT. A likely explanation for the lack of a substitution towards these airports is that, for one reason or another, they do not offer a viable alternative to nearby German airports: for example, road or rail links from Germany to these airports might not be appropriate; or the airports offer different services to German counterparts. Nonetheless, the large negative impact observed for Prague Airport is difficult to explain and is likely a consequence of some exogenous shock to passenger numbers we could not account for in our synthetic control estimations.

b) It is possible that the response of airlines, especially low-cost ones, have exacerbated the impact of AT on passenger numbers. Anecdotal evidence from the Dutch and German AT cases point to some airlines having responded to an anticipated drop in demand by relocating their services to airports outside the AT area (Zuidberg, 2015). As mentioned in Section 1.2, such responses ought to be strongest among budget airlines, since they are less tied to hubs and able to relocate quickly. Of course, the elimination of some destinations from regional airports forces travellers to shift their custom elsewhere. Our estimates are consistent with these explanations: smaller, regional airports (predominantly serving low-cost airlines) lost proportionately more passengers after the introduction of AT (see Figure 1.4). In fact, nearly all airports shown on the right side of Figure 1.3 - airports with the largest losses in passenger numbers from AT - fall under this category[22]. In this respect our findings are in line with those of Falk and Hagsten (2019).

c) Estimates for hub airports within Germany either show small negative effects (Hamburg) or positive impacts of AT (Berlin Tegel, Dusseldorf, Frankfurt Main and Munich). The greater resilience of passenger numbers at airline hubs facing AT is consistent with a lower price elasticity of demand. A likely explanation is that hubs attract a greater proportion of passengers flying on business trips, see Hess and Polak (2005); greater proportion of untaxed transfer passengers; fewer offerings from budget airlines; fewer opportunities to substitute to non-taxed routes; and a greater attachment of non-budget airlines to particular hubs (which reduces the supply side response). It is also possible that hubs within Germany gain passengers who substitute away from budget airlines once the latter airlines reduce their flights from German regional airports and from substitution induced by the relatively larger proportionate change in budget airline's ticket prices (assuming pass-on), since the

[22] Results for the two outliers, namely Berlin Schonefeld and Saarbrucken are based on ill-fitting synthetic control models.

AT due varies only by distance, not by service level.

d) While our analysis does not explicitly focus on policy outcomes, it is possible to draw some policy implications from our results. Recall from Section 1.1 that the main objectives of governments who implement ATs are: 1) to reduce the environmental burden of the aviation industry and 2) revenue raising.

From an environmental perspective, the reduction in air travel demand in Germany could act to reduce pollution levels, however the fact that some of the demand relocates to bordering airports mitigates the potential for an overall reduction in pollution levels. It is also possible that the observed changes in travel demand at hubs vs regional airports will lead to a larger share of passengers taking long-haul flights which results in greater pollution (see Mayor and Tol, 2007). The cross-border substitution effect found in our analysis also provides evidence in support of the argument that, in order to successfully curb emissions from air travel, a harmonised EU-wide aviation tax is needed (Krenek and Schratzenstaller, 2017). As long as AT rates are allowed to vary between neighbouring countries, tax avoidance through cross-border substitution will remain a salient option for mobile passengers, leading only to a relocation of air travel demand and a reduction in the intended positive environmental impact of aviation taxes (Leicester and O'Dea, 2008).

From a tax-revenue perspective, the substitution effect observed at bordering airports leads to lost tax revenues as passengers substitute to airports where they can avoid paying AT. In Section 1.4, we estimated that 4 million passengers were lost annually for German airports; assuming that half of these were departing (taxed) passengers and multiplying with the estimated average unit tax of €12.26, we estimate the annual loss in tax revenues to be approximately €24.52 million, or 2.45 percent of expected government revenues from AT (see Section 1.2). On the other hand, increased traffic at hub airports, with a higher share of business/long-haul passengers paying higher rates of AT, likely has a positive impact on average tax paid and consequently on tax revenues. Overall, the German airports included in our sample had roughly 203.9 million annual passengers combined in the post-AT years (on average); if half of these passengers were taxed, estimated tax revenues from AT amount to approximately €1.25 billion annually, slightly exceeding the policy target set by the German government.

Regardless, considering the dual objectives of the policy, in an ideal scenario AT should curb overall air travel demand (to reduce emissions), without a cross-border substitution effect that results in the loss of the tax base. Our results provide evidence to the contrary: German demand, and overall air travel demand, is not reduced to a large extent, but the cross-border effect is substantial. Please note, however, that our tax revenue estimations are based on back-of-the-envelope type calculations, as information on the number of departing passengers (the composition of the tax base) in each AT band is not available for the purposes of this research. To appropriately assess the impact of behavioural responses on policy outcomes, further analysis is needed at a more disaggregated level.

1.6 Conclusions

The synthetic control approach has provided estimates of the effects of German AT on passenger numbers using German airports and airports outside Germany but near the border. Estimates indicate that AT has significantly reduced passenger numbers, relative to the counterfactual of zero AT, for many German airports, though passenger numbers tended to hold up at and even grow somewhat at some hub airports. At the same time, most bordering airports gained passenger numbers. These findings are consistent with likely and mutually reinforcing behavioural responses of passengers and airlines to AT and the induced changes in the relative prices of airline services. From a policy perspective, these behavioural responses can have ambiguous impacts on both the environmental and revenue-raising objectives associated with AT. This chapter also provides a methodological contribution to the literature on aviation taxes: future research on aviation taxes, for example studies focusing on the ATs recently introduced in Norway and Sweden, could make use of the synthetic control approach to estimate the impact of these taxes on a variety of outcomes.

Our analysis has a number of notable limitations. In our estimations, we are unable to explicitly model the impacts of substitute forms of transportation on German (and bordering) airports[23]. For example, the deregulation of long-distance bus services in Germany in 2013 has led to a large increase in intra-urban bus traffic within the country (Knorr and Lueg-Arndt, 2016). It is possible that this had an effect on passenger numbers at some of the regional airports that offer domestic flights. In addition, we do not model tourism related factors that might drive (inbound) demand for air travel. For example, Germany reduced the VAT on hotel accommodation from 19% to 7% in 2010 (see European Commission, VAT database) which might have had a positive impact on the demand for inbound air travel. Finally, an additional limitation is that the use of aggregate data does not allow us to distinguish between the behavioural responses of different passenger types (leisure/business, short-haul/long-haul) and prohibits the estimation of tax revenues from AT. Future research, using more detailed micro-data on passengers, is needed to disentangle behavioural responses and provide an accurate assessment of tax revenue considerations.

[23] For example, there does not seem to be consistent data available on the number of bus passengers in European countries over our sample period. Eurostat has data on this for some countries but not others (German data are not available).

Chapter 2

LIMITING THE DISTORTIONARY IMPACTS OF TRANSACTION TAXES: SCOTTISH STAMP DUTY AFTER THE MIRRLEES REVIEW

2.1 Introduction

Taxes on property transactions, also known as stamp duties, are a historic and common feature of tax systems in many European countries[24]. Transaction taxes are often considered inefficient as they discourage mutually beneficial property transactions, resulting in properties not being held by the individuals who value them most (Johansson et al., 2008; Mirrlees and Adam, 2010). By disincentivising house moves at the margin, they may create distortions elsewhere, for example in the labour market by making property owners less geographically mobile, reducing labour market flexibility[25]. Consequently, the European Commission recently recommended that countries move away from transaction taxation and implement more efficient ways of taxing housing (European Commission, 2015). In a similar vein, the recent Mirrlees Review (Mirrlees et al., 2011) criticised the UK's stamp duty system and called for its abolition:

> Stamp duty has a long history in the British tax system, having first been introduced in 1694. It stems from a time when few other potential taxes were straightforward to implement, whereas the transactions on which stamp duty was levied were easy to identify and to measure. But, in the case modern case of broadly based taxation, the case for maintaining stamp duty is very week indeed (...) There is no sound case for maintaining stamp duty and we believe that it should be abolished.

Yet, despite the frequent calls for their abolition, transaction taxes remain prevalent in European tax systems due to the reliance of government budgets on the associated tax revenues (see Johannesson-Linden and Gayer, 2012). So if scrapping (or simply cutting) transaction taxes is not feasible for budgetary reasons, what policy options remain to limit

[24] See European Commission (2018) for an overview.

[25] According to Hilber and Lyytikäinen (2017) a 2 percent increase in stamp duty at the £250k threshold reduces the annual rate of mobility by 2 to 3 percentage points.

their distortionary impacts on property markets? In this chapter, we attempt to answer this question through a case study of recent transaction tax reforms in Scotland.

The Scottish Government inherited the UK-wide stamp duty regime but the Scotland Act of 2012 included stamp duty among various new devolved taxes. While abolition of stamp duty was not permissible under the Act, reform of the system to address elements of Mirrlees' critique was now an option[26]. Accordingly, in 2015, the government introduced a new Scottish transaction tax to satisfy two main policy objectives: 1) to reduce distortionary aspects of previous stamp duty regimes and 2) to make the tax system more progressive in order to encourage property market activity in the lower end of the market. This reform in Scotland resulted in several sources of variation in transaction tax rates. In our analysis, we exploit these sources of variation to investigate the diverse ways in which transaction taxes can distort property market behaviour, and also evaluate the efficacy of recent reforms in limiting these distortions. Three specific policy changes make Scotland a particularly interesting case study.

First, until December 2014, the Scottish residential property transaction tax system was characterised by a 'notched' structure – under this structure, once a given price threshold is reached, the appropriate tax rate is payable on the entire purchase price of the property, leading to discontinuous jumps in tax liability at threshold values (also referred to as price notches). Under this regime, transactions similar in terms of value incur very different tax liabilities – discouraging higher taxed transactions to a much larger extent than similarly valued but lower taxed ones, creating large incentives to transact at prices just below the threshold (Mirrlees and Adam, 2010; Best and Kleven, 2018). In December 2014, the notched structure was removed and a 'kinked' Stamp Duty Land Tax (hence-

[26] Abolishing stamp duty was not a feasible policy option for the Scottish Government. The Scotland Act 2012, which legislated devolved tax policy options for Scotland, requires stamp duty to be replaced with an equivalent transaction tax that involves interest on land (see Scottish Government, 2013).

forth SDLT) regime was introduced UK wide, including in Scotland[27]. Soon after this, the Scottish Government announced the April 1, 2015 introduction, and the precise tax schedule of a new Scottish transaction tax, the Land and Buildings Transaction Tax (henceforth LBTT). In a similar vein to the new UK SDLT, LBTT has a kinked tax schedule where only the marginal tax rate jumps discontinuously at threshold prices and the tax liability is a continuous function of price. Economic theory predicts that agents respond to both notches and kinks by bunching, that is, increased transaction activity occurs just below tax thresholds. Bunching should occur to a much larger extent in notched tax systems due to the abrupt jump in tax liability at notches (Kleven, 2016). In other words, we should expect notched tax schedules to have a more distortionary impact on property market activity than kinked ones, which was precisely the reason behind Scotland's (and the UK's) transition to a kinked transaction tax regime. Using the examples of previous and current Scottish tax regimes, we are able to investigate the distortionary impact of notches and kinks on property market activity, and assess whether replacing the notched tax regime with a kinked one was a sensible policy to reduce property market distortions.

The second margin for property market distortions we investigate in this chapter is related to the early announcement of LBTT. Early announcement created a temporary opportunity for home buyers in Scotland to save on taxes paid. LBTT has a more progressive schedule in comparison to the UK SDLT regime it replaced: that is, higher priced transactions are taxed at higher rates; with lower rates applied in some lower price ranges. As a consequence, the announcement of LBTT created a time notch in the tax schedule at April 1. Depending on the price of the property, tax liabilities were lower (or higher) if transacting before this date than after, creating large incentives for agents to strategically time their transactions where feasible. The variations in tax rates around this time notch

[27] In a kinked system, tax rates are applicable on the share of the property price above threshold values. It is worth noting that it was the Scottish Government who initially announced the planned change to a kinked system in October 2014. The UK Government applied the UK-wide change in December partly in response.

provide us with the opportunity to estimate distortions to property market behaviour in the months surrounding the policy change and determine how responsive property transactions in Scotland were to temporary savings opportunities.

Finally, the introduction of LBTT resulted in a more progressive transaction tax system in Scotland. The reform had discreetly different impacts on tax rates throughout the price distribution: effective average tax rates were 1) unaffected for the bottom segment below £125k; 2) decreased for the £125k to £380k price range; and 3) increased for the higher end of the market covering transactions over £380k[28]. The government's objective with progressive reform was to encourage transaction activity at the lower end of the market and to bring tax liabilities more in line with the ability to pay (Scottish Government, 2015). In effect, progressive reform uses tax increases in the higher end of the market to pay for tax cuts at the lower end. Ideally, the reform is intended to result in increased transaction activity for the market segment where taxes were cut; and little or no response from the market segment where tax rates increased. This way, the government achieves stimulus in the lower end of the market with little distortion to market activity (and tax revenues) in the higher end. To assess the extent to which progressive reform has accomplished these goals in Scotland, we look at the reform's impacts across the price distribution and make use of the variation in tax rate changes to investigate the permanent effect of progressive reform.

Our analysis of the property market impacts of transaction taxation makes use of a particularly rich data set containing all residential property transactions in Scotland. These data allow us to determine the price and date of each individual residential property transaction, and control for property and market characteristics. We present three main findings.

[28] Effectively, due to the short period during which the 'kinked' SDLT regime was active, and the possibility that market activity in this period was affected by the timing responses outlined above, the reform can be considered as a shift from the 'notched' SDLT to LBTT.

First, we find that the 'notched' design of the previous Scottish stamp duty system was highly distortionary to property market behaviour in that it led to the bunching of transactions below tax thresholds; the current 'kinked' system still results in bunching, but this is small and only occurs at a subset of thresholds. Second, we find that the temporary tax savings opportunity created by the announcement of future tax reform did lead to a substantial behavioural response as many home buyers changed the timing of their transactions to save on the associated tax liabilities. Finally, our evidence suggests that progressive transaction tax reform in Scotland led to a substantial increase in transaction activity in the lower end of the market, where tax rates decreased; had no discernible effect on most price ranges where tax rates increased; and only had a significant negative effect on the market for very expensive properties. From a policy perspective, our findings suggest that the Scottish Government's objectives with respect to reforming the stamp duty system were accomplished: the shift to a kinked tax schedule minimised distortions from bunching responses, while progressive reform was successful in encouraging market activity in the lower end of the market (without distorting activity in the higher end to a large extent). Nonetheless, significant distortions persist due to the continued re-emergence of time notches upon announcement of new policies.

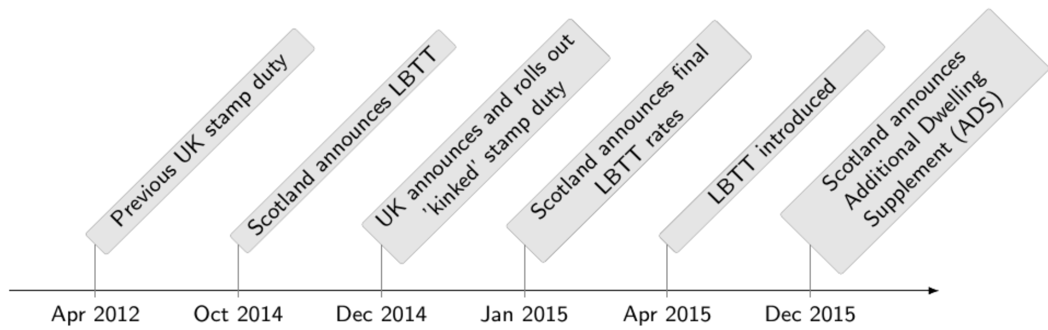
This chapter makes several contributions to the empirical literature on the effects of transaction taxes on property market behaviour. First, while a number of recent studies find evidence of distortions created by notched transaction tax regimes (Kopczuk and Munroe, 2015; Best and Kleven, 2018; Slemrod et al., 2017) our study is the first one to look at the distortionary effects of kinks, and investigate the transition from notched to kinked tax schedules. We also contribute to the literature on the effects of time notches. Prior case studies find evidence of timing responses to the pre-announced end date of temporary fiscal stimulus (Best and Kleven, 2018); and preceding changes in flat tax rates (Fritzsche and Vandrei, 2019). Our study is the first one to look at the effects of time notches across the price distribution and especially on the market for expensive properties. Lastly, we contribute to the literature on the permanent effects of transaction taxation (see Besley et al., 2014; Slemrod et al., 2017; Best and Kleven, 2018; Fritzsche and Vandrei, 2019) through an analysis of progressive transaction tax reform in Scotland. Uniquely in the literature, our case study allows us to look at the property market impacts of progressive transaction tax reform, and assess the heterogeneity of impacts across the price distribution. Moreover, previous UK case studies on the effects of transaction taxation use variation in tax rates from temporary tax cuts for cheaper properties during periods of economic downturns (Besley et al., 2014; Best and Kleven, 2018). These studies are therefore mostly focused on the effects of fiscal stimulus on a set of lower priced properties, while our study focuses on the effects and policy implications of comprehensive tax reform throughout the price distribution.

The remainder of the chapter is organised as follows. Section 2.2 provides the relevant policy background. Section 2.3 reviews the related literature. Section 2.4 describes the data. Section 2.5 outlines our empirical approach and summarises our results. Section 2.6 provides a discussion of our findings. Section 2.7 concludes.

2.2 Background

Tax on property transactions, also known as stamp duty, is paid by buyers of properties every time a new transaction takes place. Transaction tax rates are different for residential and commercial transactions. In our analysis, we consider the residential property market and the corresponding tax system. In Scotland, the residential transaction tax system has gone through several changes in recent years. We consider as the starting period of our analysis the previous UK stamp duty system as at April 2012[29]. The timeline of relevant changes to the Scottish property transaction tax system is presented in Figure 2.1.

Figure 2.1: *Timeline of recent changes to the Scottish residential property transaction tax system*



In Scotland, the UK stamp duty system (known as Stamp Duty Land Tax, or SDLT) was in place until the introduction of the Land and Buildings Transaction Tax (LBTT) in April 2015. Initially, Scotland announced the introduction of LBTT in October 2014[30] However, in December 2014, the UK government changed the SDLT system from the previous 'notched' to a 'kinked' regime and altered the rates and threshold values. This tax regime continued to apply to Scotland until April 2015. After the UK wide change in the stamp duty system, the Scottish Government announced, in January 2015, final LBTT rates and thresholds to be in place from April 1, 2015. Later in 2015, they also announced the introduction of the Additional Dwelling Supplement (ADS), a three-percent surcharge on second properties priced over £40,000[31]. In effect, three transaction tax regimes were in place in Scotland within our sample period of April 2012 to December 2017:

[29] This period also corresponds to the end of a two-year period of first-time buyers' relief. The stamp duty system had different bands before, and we would like to avoid bias from this.

[30] The initial announcement has already led to the anticipation of a new tax regime (to be implemented in April), even though final LBTT rates were not known at the time. Awareness of future tax changes may have already encouraged tax planning behaviour from as early as October 2014.

[31] Unfortunately, we are unable to explicitly model the effects of ADS as it is not possible to identify second properties in our data set.

1. The previous (UK wide) SDLT regime with the notched structure from April 2012 to December 2014

2. The new (UK wide) SDLT regime with the kinked structure from December 2014 to March 2015

3. The LBTT regime, also following a kinked structure but more progressive than UK SDLT, from April 2015

Nominal and effective average tax rates under each regime are summarised in Figure 2.2 and Figure 2.3. Key features and changes in recent transaction tax regimes in Scotland emerge from these comparisons.

First, the notched structure of the previous SDLT regime resulted in large jumps in both effective average and marginal tax rates at threshold values (price notches). These price notches created incentives to transact at prices just below threshold values. After tax reform, UK SDLT and Scottish LBTT both follow a kinked tax schedule. In a kinked schedule, only the marginal tax rate jumps discontinuously at thresholds while tax liability remains a smooth function of property price (see the solid and dashed lines in Figure 2.3). In Scotland, the shift towards the kinked tax regime was justified as a measure to create a more sensible tax system, in which transactions similar in terms of value are taxed to a similar extent (Scottish Government, 2015).

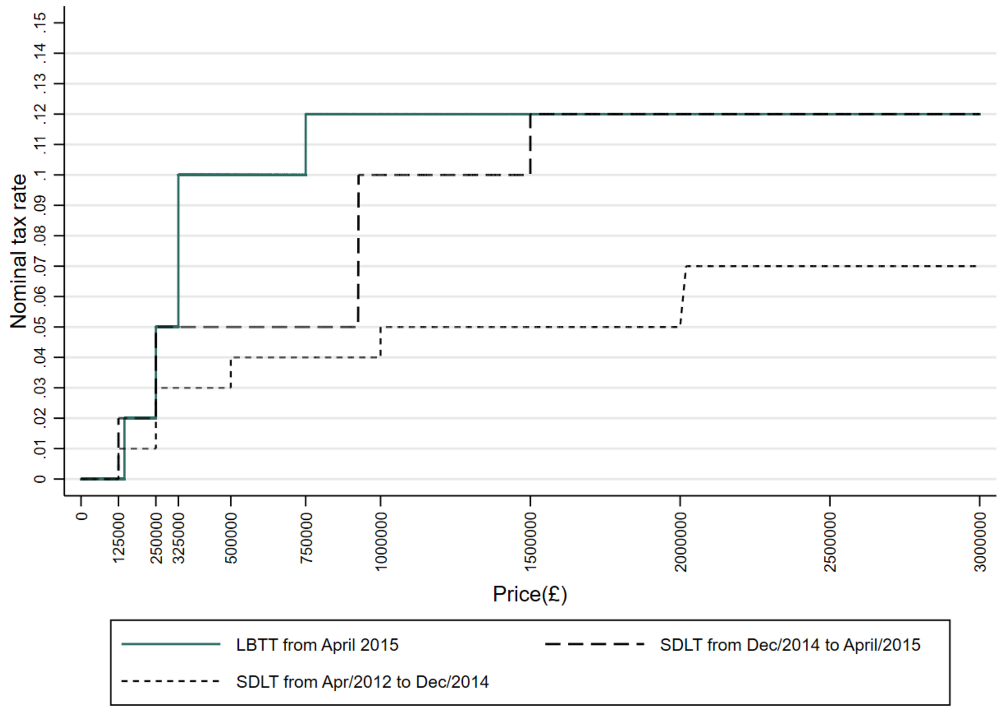
Second, the introduction of LBTT created a time notch in the Scottish transaction tax system at April 1, 2015. In Figure 2.3, the area between the dashed and solid lines highlights the differences in tax rates payable before and after this date. Being taxed under SDLT before April 1, buyers of properties in Scotland had the incentive to change the timing of transactions and be taxed under the more generous tax regime. In the price range £125k to £333k, there was an incentive to delay transactions; while in the over £333k price range there was an incentive to bring transactions forward[32].

Lastly, the introduction of LBTT in Scotland has made the transaction tax regime in Scotland more progressive – with lower tax rates at low prices, and higher tax rates at high prices. The objective of progressive reform was to boost transaction activity in the lower end of the market and make the tax burden more in line with the ability to pay (Scottish Government, 2015). As the 'new' SDLT system was only in place during the announcement period of LBTT, property market activity in this period was likely influenced by the timing responses outlined above[33]. Therefore, the tax reform in Scotland is effectively a shift from the previous (notched) stamp duty system to LBTT. The reform led to lower effective average tax rates between £125k and £380k and higher effective average tax rates over £380k (see Figure 2.3).

[32] Albeit a tiny difference in effective average tax rates, this is visible in Figure 2.3 as the dashed line is slightly left to the solid line for the £125k to £ 333k price range.

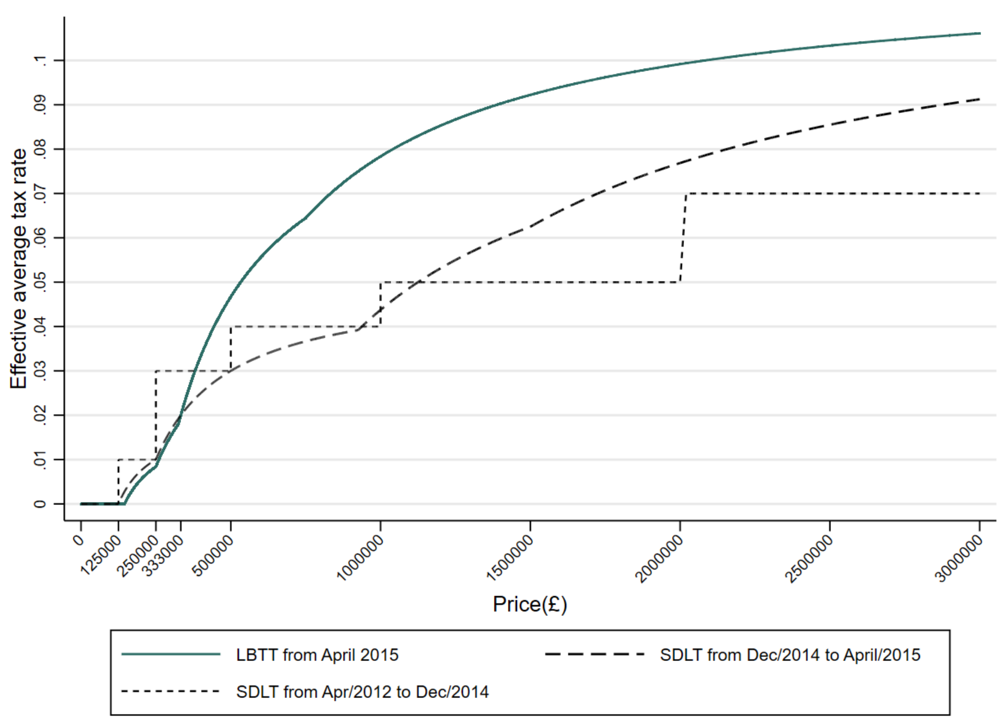
[33] Indeed, we find overwhelming evidence of this in our empirical analysis below.

Figure 2.2: *Nominal tax rates in the Scottish transaction tax system under different tax regimes*



Notes: The numbers on the horizontal axis correspond to price thresholds where nominal tax rates change under each tax schedule.

Figure 2.3: *Effective average tax rates in the Scottish transaction tax system under SDLT and LBTT tax regimes*



Notes: The numbers on the horizontal axis correspond to price thresholds where effective average tax rates change under each tax schedule. The threshold of 333k indicates the price over which effective average tax rates under LBTT rise above those under UK SDLT.

2.3 Related literature

This chapter contributes to the burgeoning literature on the impacts of transaction taxes on property markets. Several studies (Kopczuk and Munroe, 2015; Slemrod et al., 2017; Best and Kleven, 2018) investigate the distortions created by price notches in transaction tax schedules. All these studies find significant evidence of bunching of transactions below price notches, although magnitudes vary according to the specific property market setting. In our analysis of bunching around price notches in Scotland, we employ the bunching estimator methodology. This methodology was originally outlined in Chetty et al. (2011) and is also used in Best and Kleven (2018) who analyse the impact of the 2008/09 stamp duty holiday on the UK property market[34]. This approach allows us to estimate counterfactual transaction densities for the price distribution – how transaction densities would be in the absence of price notches. To assess the impact of price notches, we can then compare actual densities to counterfactual ones in the regions of notches. Our study is unique in the literature in that it does not limit its focus to bunching near price notches, but also looks at similar responses around kink points. Using the current Scottish transaction tax system as a case study, we find some evidence of bunching around kinks, though these responses are far less substantial than similar responses to price notches. This finding is in line with theoretical results that predict limited bunching (in comparison to notched schedules) in kinked tax schedules (Saez, 2010; Kleven, 2016).

A few studies also investigate the impact time notches in transaction tax regimes have on the timing of transactions. For example, Slemrod et al. (2017) investigate the impacts of simultaneous price and time notches in the Washington D.C. transaction tax system and find no evidence of agents timing transactions in response to the time notch. The authors suggest that the null effect observed might have been a consequence of the costs of

[34] The stamp duty holiday was a temporary abolition of transaction taxes in the UK for the £125k to £175k price range.

optimisation frictions outweighing the benefits from timing transactions for most home buyers. Conversely, using a panel regression approach in which they regress changes in transaction activity on dummy variables indicating months before (and after) tax changes, Fritzsche and Vandrei (2019) find evidence from Germany that in months before transaction tax hikes, transaction volumes for single-family homes increased by roughly 40%. Finally, Best and Kleven (2018) find that the anticipated ending of the stamp duty holiday led to significant bunching of transactions just before the end date. In other words, anticipating the end of the tax-exempt period, buyers of properties in the UK brought transactions forward to realise tax savings.

In our analysis of time notches, we build on the methodology used in Best and Kleven (2018). Our estimation relies on comparing actual time trends in transaction activity (in the months near the time notch) to counterfactual time trends. In Best and Kleven (2018) these counterfactual time trends are constructed using the price ranges unaffected by the time notch – differences in the bunching of transactions near the time notch between affected and unaffected price ranges are indicative of timing responses. They motivate this approach as means to control for the time notch being near Christmas – transaction activity is generally non-existent this time of the year leading to idiosyncrasies in market behaviour. In our case, the time notch is during the spring period when transaction activity is 'normal', and we only need to control for seasonality. We therefore use a simpler bunching estimator strategy than the one outlined in Best and Kleven (2018): our counterfactual time trends are based on predicted values from a fixed-effects time-series regression that models transaction activity using seasonal and cyclical time trends.

Lastly, there are attempts in the literature to estimate the permanent effect increased transaction tax rates have on property market activity. This effect, also referred to as the 'lock-in' effect, arises because increased property transaction taxes make it more costly to transact, having a lasting impact on how often properties are bought and sold (Slemrod et al., 2017). Studies attempting to estimate the permanent effect of transaction taxes often use price, time, or geographical notches in tax systems to identify the long-term effect of changing tax rates (Dachis et al., 2011; Best and Kleven, 2018). The estimates from these studies are however local to the regions of the notches investigated. Other studies (Besley et al., 2014; Slemrod et al., 2017) use a difference-in-differences (diff-in-diff) approach where they compare price ranges where tax rates changed to price ranges where they did not. Best and Kleven (2018) use a similar approach to examine medium-term timing and extensive margin responses to the 2008/09 stamp duty holiday. An advantage of this approach is that it allows for a convenient comparison group when only certain price ranges are affected by policy change (and others are not). Nonetheless, applying diff-in-diff in this context comes with two notable threats to identification: 1) affected and unaffected price ranges might follow different time trends before the policy change (absence of parallel trends) and 2) endogenous selection into the treatment group, as a consequence of, for example, agents moving into a particular price range to qualify for tax breaks (see Best and Kleven, 2018).

In our empirical analysis below, we follow previous studies and employ a diff-in-diff approach to examine the permanent effects of progressive LBTT reform through a comparison of, over time, price ranges affected (treated) and unaffected (control) by stamp duty reform in Scotland. We validate the parallel trends assumption by plotting time trends and using event study plots. To avoid bias from endogenous selection into treated price ranges, we perform a robustness check where we exclude price regions near threshold prices. Uniquely in the literature, we also disaggregate our results by narrow price ranges to examine how progressive reform has affected different parts of the price distribution.

2.4 Data

To analyse the impact of transaction taxes on the Scottish property market we use a data set containing the universe of Scottish property transactions from the Registers of Scotland (ROS)[35]. ROS are a non-ministerial government department responsible for compiling public registers in Scotland. Our data set includes all property transactions in Scotland during the sample period (April 2012 to December 2017). It also contains the types, addresses, postcodes, coordinates, and postcode areas of properties, along with, most importantly, the date and price at which they were sold. As our analysis focuses on the residential property market, we restrict our sample to include only residential transactions. Furthermore, we exclude transactions for which no price was given, or for which nonsensical values were given[36]. This latter group of transactions constitutes roughly 0.4% of our raw data. We observe when a property is transacted through the variable 'date of entry' which notes the date at which ownership rights were exchanged between the buyer and the seller. In our sample period, we have 716,275 property transactions in Scotland, and mean transaction price is £165,073.4.

[35] Access to this data was granted through the Glasgow Urban Big Data Centre (UBDC).

[36] Possible reasons for 'no price' transactions include ownership changes due to divorce settlements or inheritance.

2.5 Identification strategy and results

Our empirical investigation concerns the following margins for distortions in recent Scottish transaction tax regimes: 1) price notches in the previous stamp duty system 2) kinks in the LBTT system 3) the time notch at the April 1 introduction of LBTT and 4) the shift from the previous stamp duty system to the more progressive LBTT system.

In this section, we estimate how responsive property market activity in Scotland was to changing tax rates along each of these margins. While we employ distinct identification strategies for each case, we also control for biases arising from their combined effect on the property market.

2.5.1 Estimating the property market impacts of notches and kinks in Scottish transaction tax regimes

2.5.1.1 Estimating the impacts of price notches in the previous SDLT system

The previous (UK wide) SDLT system was in place between April 2012 and December 2014 (see Section 2.2)[37]. Transaction taxes were charged on the entire purchase price of the property creating 'price notches' in the tax schedule. At price notches, threshold prices

[37] As noted in Section 2.2, the beginning of this period corresponds to the end of the period for first-time buyer's relief. The anticipated end of the relief period may have encouraged agents to transact before the new tax regime was implemented, potentially resulting in a number of missing transactions in the months after April 2012. This is unlikely to bias our empirical analysis below as we aggregate transactions over the entire previous stamp duty period, meaning the timing effect in the first month or two will likely be negligible on aggregate. Nonetheless, as a robustness check, we performed our analysis below excluding the first three months of the previous stamp duty period from the sample. Dropping these months had no discernible impact on our results.

where average tax rates change, both the marginal tax rate and tax liability jump discontinuously creating large incentives for buyers to transact at prices just below a given tax threshold and avoid the abrupt increase in taxes payable. These incentives may lead to the bunching of transactions at prices just below notches. Bunching responses might be accentuated by the fact that notches create particularly salient reference points for agents to strive towards (Kleven, 2016).

Tax rates and thresholds for the previous SDLT system are summarised in Table 2.1, along with the jump in tax liability, corresponding to a £1 increase in the transaction price, at each price notch. It is evident from the numbers presented here that the Scottish system has created particularly large and salient incentives to reduce transaction prices in the regions of notches, as doing so even by a small amount could lead to a monetary saving of several thousand pounds[38].

Table 2.1: *Previous UK Stamp Duty (SDLT) - Tax Schedule and Price Notches*

Price	**Tax rate**	**Jump in liability at threshold**
£0 - £125k	0%	£0
£125k - £250k	1%	£1,250
£250k - £500k	3%	£5,000
£500k - £1m	4%	£5,000
£1m - £2m	5%	£10,000
Over £2m	7%	£40,000

In Figure 2.4 and Figure 2.5 we plot the density of property transactions at different prices. The figures aggregate transactions in Scotland for the entire period during which the previous stamp duty system was in place. These plots already provide us with visible in-

[38] The size of the monetary tax savings is important because optimisation frictions, such as the cost of renegotiating the transaction price, might dissuade buyers from changing their behaviour if the realised tax saving from doing so is small (Kleven, 2016; Slemrod et al., 2017).

dications of bunching, that is, a higher density of property transactions just below price notches (and lower density just above). We can observe spikes in the density of transactions corresponding to the £125k, £250k, and £500k price notches and there is also some bunching around the £1m notch though this does not seem to be outstanding relative to the price neighbourhood. Note however, that bunching can occur for reasons other than the incentives created by notches, for example by salient reference points at round number prices (Kleven, 2016). To disentangle bunching responses to price notches, we therefore need to control for other factors that could influence the density of transactions throughout the price distribution. We do this using the bunching estimator approach outlined in Chetty et al. (2011). This approach relies on the estimation of a counterfactual density of transactions – what the density of transactions would look like in the absence of price notches at threshold values. As noted in Section 2.3, our use of these estimators follows Best and Kleven (2018). We estimate counterfactual transaction densities by fitting a 5-th degree polynomial to the data, using the wider region of the price notch but excluding observations in a small range around the notch.

More specifically, the counterfactual distribution is estimated using the following regression model:

$$t_i = \gamma_0 + \sum_{j=0}^{5} \gamma_{1j} p_i^j + \sum_{r=1}^{3} \gamma_{2r} R_{ri} + \mu_i \tag{2.1}$$

where t_i are the number of transactions in price bin $'i'$ and price bins are £5,000 wide[39]. The second term on the right-hand side is a 5-th degree polynomial of the transaction price $'p'$ aimed at approximating the relationship between an increase in price and the density of property transactions. The third term is a set of indicator variables for transactions at round number prices for multiples of £10,000, £25,000, and £50,000. This is to control for possible bunching at round prices, and transactions being more fre-

[39] In similar studies such as Best and Kleven (2018) or Slemrod et al. (2017) £100 (or $100) price bins are used. However, in the case of the Scottish property market, if aggregated at that level most price bins contain zero transactions as most transactions tend to occur at large round numbers. For this reason, we decide to aggregate at the £ 5,000 level.

Figure 2.4: *Density of transactions at different prices in Scotland – previous stamp duty (SDLT) system from April 2012 to December 2014*

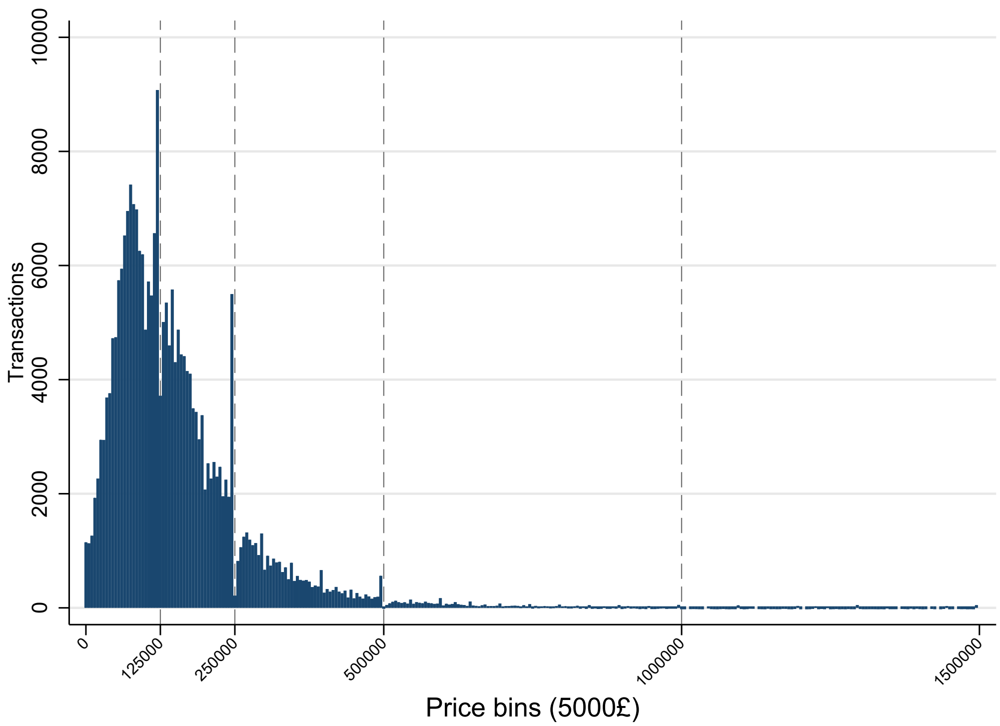
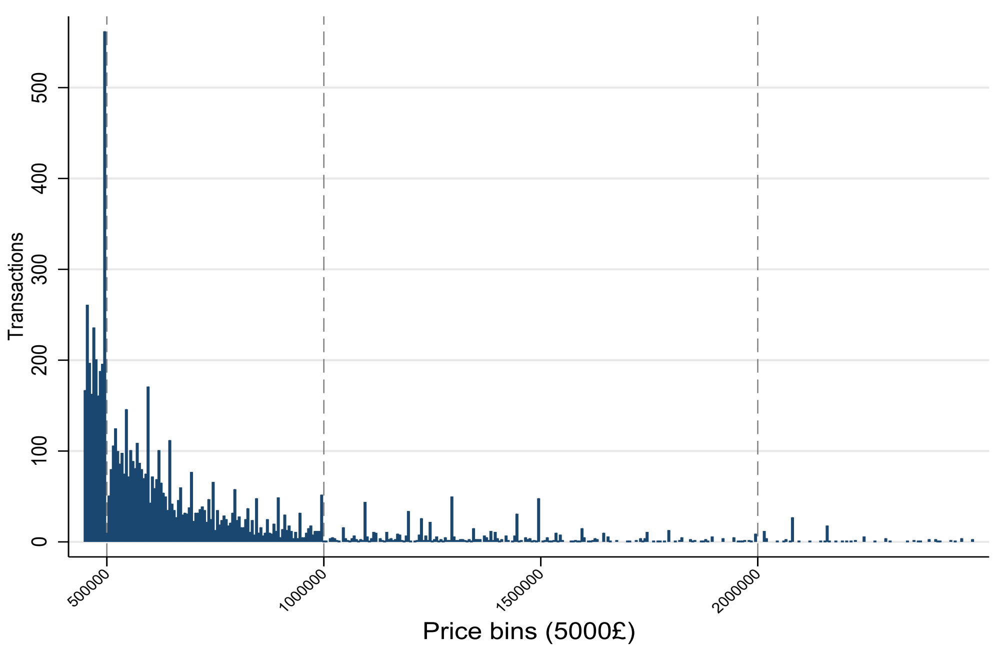


Figure 2.5: *Density of transactions at different prices in Scotland (higher price range) - previous stamp duty (SDLT) system from April 2012 to December 2014*



Notes: Vertical dashed lines indicate threshold values where tax rates change (price notches).

quent at some round prices than others[40]. We exclude all transactions from the [-£10,000; +£10,000] range around the price notch so that bunching around the notch in the actual data does not affect our counterfactual distribution[41]. To assess whether the density of transactions increased below the price notch (and dropped above), we plot the predicted (counterfactual) transaction density for each price bin and compare it to the actual transaction density. The plots are summarised in Figure 2.6[42]. The vertical difference in transaction density between actual (solid line) and counterfactual (dashed line) plots in the region of the price notch should be indicative of the presence and extent of bunching responses. In our case, two estimates are of interest: 1) the excess mass, or increased transaction density relative to the counterfactual, just below the price notch and 2) the missing mass, or reduced transaction density, just above the notch. These estimates are presented in Table 2.2.

There are two limitations of this analysis that are worth noting here. First, the bunching estimator procedure only yields a 'partial counterfactual' as extensive margin responses (to buy or not to buy a property) are not incorporated (Kleven, 2016). Second, estimation of the number of missing transactions (missing mass) above price notches is likely to be imprecise as it is highly sensitive to parametric assumptions (Kleven and Waseem, 2013; Kopczuk and Munroe, 2015). Similarly to Best and Kleven (2018), we can largely ignore these concerns based on the size, significance and robustness of the bunching response observed near price notches (see below). This is because inference in our case is predicated upon finding evidence of bunching *per se* and does not rely on precise estimates of the components of bunching responses.

[40] We estimate round number bunching using the segments of the empirical distribution that are not near price notches (not in the excluded range).

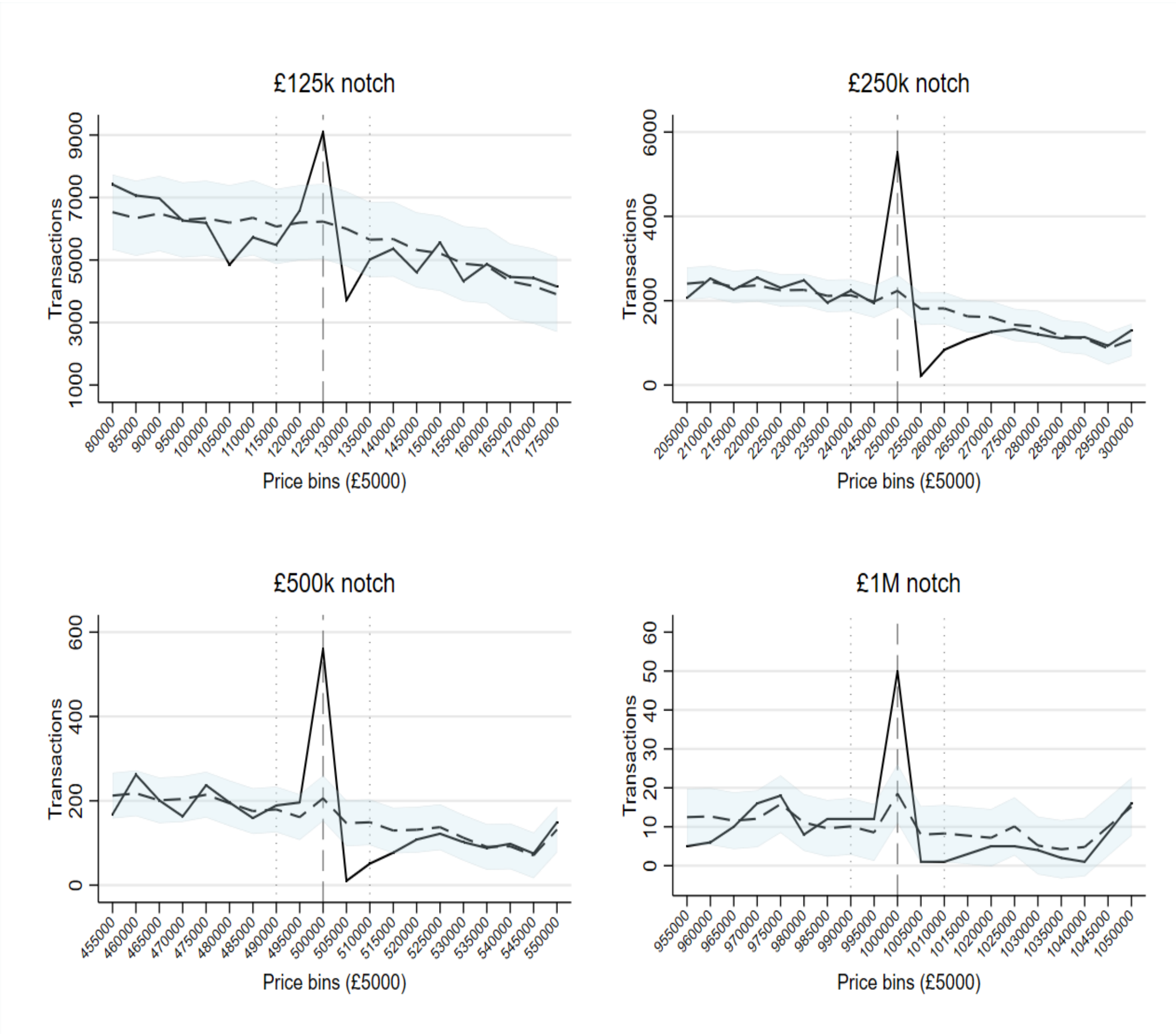
[41] Our estimates our robust to changing the size of the excluded region by £5,000 on either side, and are also robust to the order of the polynomial used in the model.

[42] In the neighbourhood of the £ 2m notch, there are very few transactions and the density follows an idiosyncratic pattern. We therefore cannot rely on our counterfactual estimation and will not analyse bunching around this notch.

Except for the £125k price notch analysis, our counterfactual estimates fit actual data relatively well in the price regions not near the price notches. For all price notches, we find strong evidence of bunching behaviour – actual densities below price notches are significantly (as indicated by the 95% confidence interval around our counterfactual estimates) higher than counterfactual predictions in all cases. In the £5,000 price bin just under the price notch we estimate that, in comparison to the counterfactual, transaction volumes are approximately 1.5, 2.5, 2.7, and 2.7 times larger for the notches at £125k, £250k, £500k, and £1m, respectively. We can also note that in all our estimations the higher density of transactions under the notch is followed by a lower density of transactions at prices just above (see Table 2.2). The estimates summarised in Table 2.2 indicate that the missing mass above price notches tends to be larger than the excess bunching just below, although in most cases these differences are not statistically significant. According to Kopczuk and Munroe (2015), the larger missing mass relative to excess bunching may be due to extensive margin responses, i.e. buyers and sellers not transacting at all in the region of the notch. Looking at Figure 2.6, we can also note that the missing mass spans a £15,000 - £20,000 range above each price notch, suggesting that some agents were willing to reduce transaction price by much more than the increase in tax liability at the threshold (see Table 2.1). In their theoretical model, Best and Kleven (2018) stipulate that this phenomenon is likely a result of downpayment constrained home buyers, with high loan-to-value ratios on their mortgages, being particularly responsive to transaction tax changes. These home buyers are likely to be more sensitive to tax changes because they need to pay transaction taxes upfront and cannot package these taxes into mortgage loans.

Overall, our findings are indicative of agents manipulating prices to fall just under the relevant tax threshold. Our results show that the Scottish property market was highly responsive to the presence of price notches under the previous SDLT system. According to our estimates, the overall tax revenue leakage (counterfactual tax revenues minus actual tax revenues) in the price regions included is approximately £15.7 million, or roughly 5.3% of overall counterfactual tax revenues[43].

Figure 2.6: *Bunching at price notches in the previous SDLT system*



Notes: Solid lines represent the actual density of transactions at different prices. Dashed lines represent the counterfactual density of transactions, calculated using predicted values from the regression model outlined above. The blue shaded area around the counterfactual represents the 95% confidence interval of the counterfactual estimates. The actual distribution spanning the confidence intervals is indicative of a significant bunching (or missing mass) estimate.

[43] This is only a 'naive' estimate as extensive margin responses to the notch are not incorporated in the counterfactual.

Table 2.2: *Bunching at price notches under the previous SDLT system*

Price notch	*Excess bunching*	*Missing mass*
£125k	1.460** (0.609)	1.617** (0.638)
£250k	2.466*** (0.598)	3.324*** (0.652)
£500k	2.720*** (0.644)	3.395*** (0.864)
£1m	2.699*** (0.718)	3.610 (2.846)

Notes: Excess bunching is estimated from the £5,000 price bin just below a given notch and is relative to the counterfactual density in the same price range. The missing mass is calculated from the excluded range above the notch and is also relative to the average counterfactual density in this price region. Standard errors (in parentheses) are obtained through a bootstrapping procedure (see Best and Kleven, 2018). *** p<0.01, ** p<0.05, * p<0.1.

2.5.1.2 Estimating the impacts of kinks in the LBTT system

By changing the way in which tax liabilities are calculated at different prices, the LBTT reform effectively replaced the notches in the Scottish transaction tax schedule with convex kink points[44]. Contrary to notches, at (convex) kink points only the marginal tax rate jumps discontinuously while tax liability is a continuous function of price. According to Kleven (2016), kinked tax schedules lead to a change in the slope of agents' budget line at kink points. Agents in the region just above kink points will be able to attain a higher indifference curve (tangent to the pre-kink budget line) and will move below the kink, leading to bunching. Convex kinks should not however lead to a missing mass above kink points as agents move down the new budget line in response to higher marginal tax rates in order to fill up the 'hole'. Consequently, in a kinked transaction tax schedule we should observe bunching at prices just below tax thresholds, but no missing mass at prices just above. In order to assess whether this is the case under the current Scottish transaction tax system, we use the bunching approach from the previous section to estimate coun-

[44] Note, that while the UK-wide reform to the SDLT system has already replaced the notches in December 2014 (see Section 2.2), we do not consider this period in our bunching analysis as transaction activity was highly affected by timing responses (see Section 2.5.2 below) during the few months the new UK SDLT was in place in Scotland.

terfactual transaction densities near the kink points in the LBTT schedule[45]. LBTT rates change at threshold values of £145k, £250k, £325k and £750k, corresponding to nominal tax rates of 2%, 5%, 10% and 12%, respectively (see Section 2.2, Figure 2.1). Bunching estimates for the price regions near these kink points are summarised in Figure 2.7; corresponding estimates of excess bunching and missing mass are shown in Table 2.3.

As shown in Figure 2.7, bunching responses under the kinked LBTT system are much less substantial in comparison to those under the notched schedule of the previous SDLT. In fact, we find no evidence of bunching for the £145k and £325k kink points; a small bunching response at the £250k kink point; and a more substantial bunching response at the top rate threshold of £750k. Our estimates for the missing mass of transactions above kink points are small and not significant at the 5% level (see Table 2.3).

The finding of bunching responses below kink points, and no missing mass above them, are consistent with the stipulations of the theoretical literature, particularly for kink points at round numbers that provide salient reference points for agents to strive towards (Saez, 2010; Kleven, 2016). Note however, that our counterfactuals predict only minimal round number bunching, especially in comparison to the notched case above, at the threshold prices of £250k and £750k (see Figure 2.7). This is likely a consequence of the empirical distribution being smooth outside the price regions near kink points (this is the price region we use to estimate round number bunching). Since we observe no discernible bunching response at the $145k and £325k kinks, it is possible that the bunching observed at other kink points is more a consequence of agents responding to particularly salient reference points than a response to a change in the marginal tax rate.

[45] The data used for this analysis are from the period June 2015 to December 2017. The first two months of the LBTT period were excluded to avoid the timing responses to the new tax regime interfering with our results.

Figure 2.7: *Bunching at kink points under the LBTT system*

Notes: Solid lines represent the actual density of transactions at different prices. Dashed lines represent the counterfactual density of transactions, calculated using predicted values from the regression model outlined above. The red shaded area around the counterfactual represents the 95% confidence interval of the counterfactual estimates. The actual distribution spanning the confidence intervals is indicative of a significant bunching (or missing mass) estimate.

Table 2.3: *Bunching estimates for kink points in the LBTT system*

Kink point	*Excess bunching*	*Missing mass*
£145k	0.925 (0.582)	1.004* (0.594)
£250k	1.223** (0.585)	1.174* (0.618)
£325k	0.841 (0.626)	1.126 (0.697)
£750k	2.125*** (0.698)	1.131 (1.085)

Notes: Excess bunching is estimated from the £5,000 price bin just below a given kink point and is relative to the counterfactual density in the same price range. The missing mass is calculated from the excluded range above the kink point and is also relative to the average counterfactual density in this price region. Standard errors (in parentheses) are obtained through a bootstrapping procedure. *** p<0.01, ** p<0.05, * p<0.1.

2.5.2 Estimating the impact of a time notch at the end of the LBTT announcement period

As previously explained in Section 2.2, the UK government removed the 'notched' structure of the stamp duty system and introduced a new tax regime for SDLT in December 2014. This new tax regime applied to Scottish property transactions as well, albeit only temporarily. In January 2015, the Scottish Government announced the April 1 introduction and specific tax schedule of LBTT, a more progressive tax with higher tax rates at higher property prices. The announcement period, which lasted from 21 January 2015 until the end of March 2015, provided buyers of properties in Scotland with a temporary savings opportunity[46]. Those aiming to transact in the £125k to £333k price range, where tax liabilities are lower under LBTT than under the new SDLT system, had the incentive to delay transactions and wait to be taxed under LBTT. On the other hand, in the over £333k range, buyers had the incentive to bring transactions forward and pay lower taxes under SDLT.

[46] Also, the initial (October 2014) announcement (see Section 2.2) of LBTT might have already led to tax planning behaviour from agents anticipating a tax change in April.

Figure 2.8 summarises the temporary savings opportunities available to buyers during the announcement period. Note, that while savings from delaying transactions in the £125k-£333k price range are relatively low, savings from bringing transactions forward in the higher price range can be almost 4% of the property value. On the other hand, delaying transactions might be easier than bringing them forward given the time it takes to move from offer to ownership on the property market[47]. Based on this, we can form the following expectations:

1. *In the £125k to £333k price range*: lower than average transaction volumes in the months preceding April 2015 due to buyers delaying transactions, and higher than average transaction volumes during and after April 2015 due to those transactions taking place under LBTT. As less time is required for sellers to 'hold' properties, delaying transactions should become decreasingly costly as we move towards April, meaning the decline in transaction activity should be gradual in the preceding months. After April 1, transaction activity should increase sharply to return to normal levels as there is no real incentive to delay transactions any further.

2. *In the over £333k price range*: higher than average transaction volumes in the months preceding April 2015 due to buyers bringing transactions forward, and lower than average transaction volumes during and after April 2015 due to those transactions not taking place under LBTT. The timing response should be strongest just before April, as there is no incentive (and potentially no time) to bring transactions forward to many weeks before this. After April 1, we should then expect a gradual decline in transaction numbers as transactions predicted for the next months no longer take place.

[47] According to Besley et al. (2014) the time from first offer to change in ownership for UK properties is usually under 60 days.

Figure 2.9 and Figure 2.10 show monthly transaction volumes in the £125-£333k and over £333k price ranges, respectively. While there is a clear drop in transactions in the months before April 2015 in the lower price range, property market activity is generally low at this time of the year and it is difficult to disentangle seasonal effects from the impact of the behavioural response. On the other hand, in the over £333k price range (see Figure 2.10), bunching around the time notch is evident from the unusually high number of transactions in March 2015.

Figure 2.8: *Gains and losses from being taxed under the SDLT regime as a share of the property value (in comparison to being taxed under LBTT)*

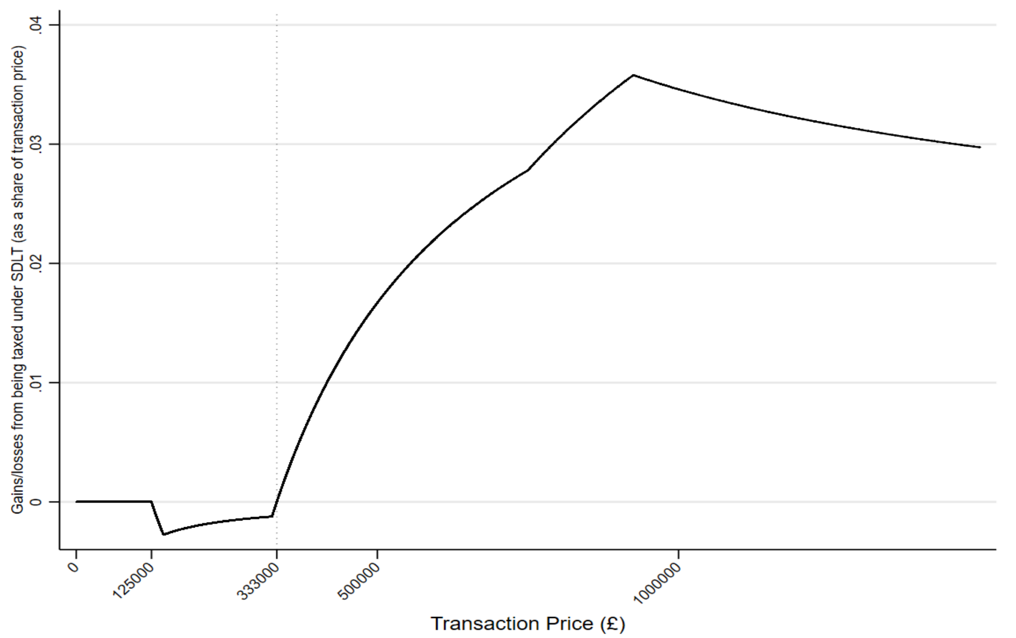


Figure 2.9: *Monthly transaction volumes between £125k and £333k*

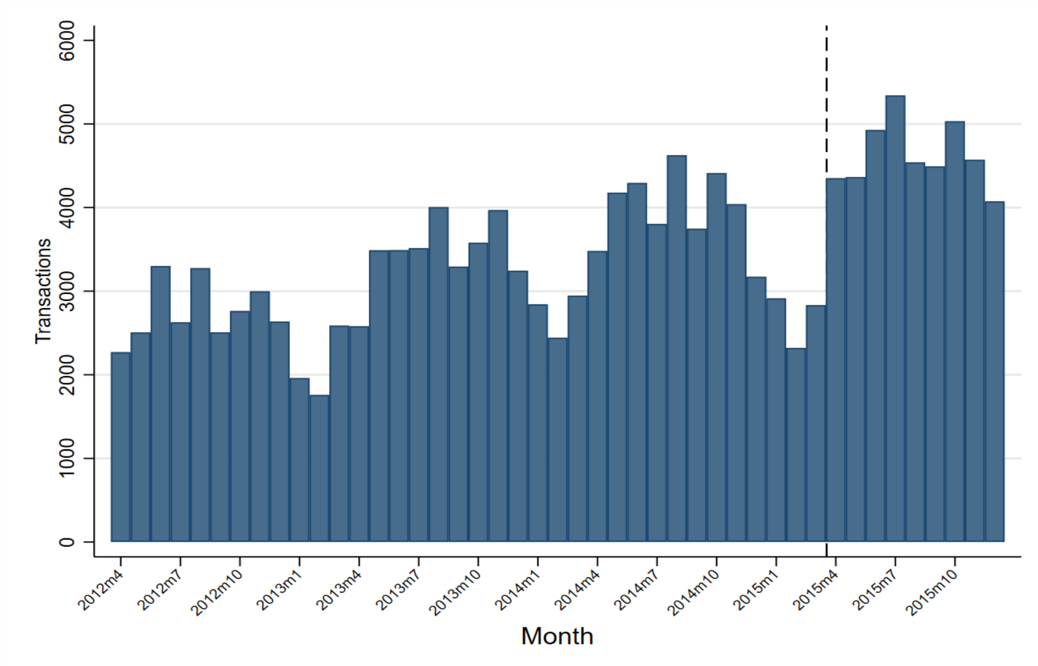
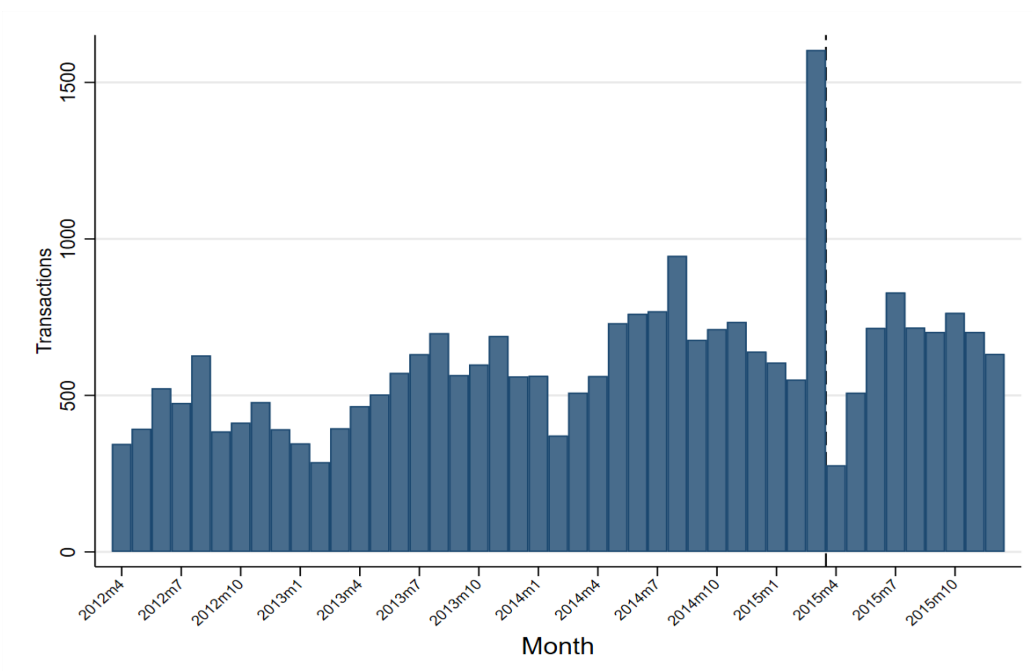


Figure 2.10: *Monthly transaction volumes over £333k*



Notes: The vertical dashed line indicates the introduction of LBTT on April 1, 2015.

Our empirical strategy aims to isolate the timing responses arising from temporary savings opportunities from the daily, monthly, and yearly time trends affecting the property market. Analogously to our approach in Section 2.5.1, we estimate counterfactual monthly distributions of property transactions over time. We do this by fitting a fixed-effects time-series model to the empirical data, excluding the months around the April 2015 LBTT introduction that might be affected by reduced/increased transaction volumes due to timing responses. This counterfactual time trend approximates a scenario where no tax savings opportunities are present during the LBTT announcement period. We then compare the counterfactual time trends to actual time trends to estimate the impact of behavioural responses on transaction numbers in the months around April 2015. Our counterfactual is based on the following fixed-effect regression model:

$$t_{im} = \gamma_0 + \theta_y + \theta_m + \theta_{dow} + \theta_{dom} + \theta_p + \epsilon_{im} \tag{2.2}$$

where t_{im} are transaction volumes in price bin $'i'$ and month $'m'$ and $\theta_y,\ \theta_m,\ \theta_{dow}, \theta_{dom},\ \theta_p$ are fixed-effects for year, month, day of the week, day of the month, and price bin, respectively[48]. From this estimation we exclude the announcement period (December 2014 to April 2015) and the three months following the introduction of LBTT[49]. Our sample period covers the previous SDLT system (from April 2012) and the LBTT system until the announcement of the Additional Dwelling Supplement (ADS), a tax on second properties, in December 2015[50]. Our model is estimated for both the £145k to £333k, and the over £333k price range[51]. We plot the counterfactual time trends estimated using Equation 2.2 against the actual time trends in transaction volumes. We associate the vertical difference

[48] We include price range fixed effects to control for the presence of notches near some price ranges in the early part of the sample period.

[49] Our results do not change when four months are excluded.

[50] The ADS was announced during the last days of November, however, since it was only introduced next April, we assume no behavioural responses until 2016 (there would not be enough time to file and process property transactions in 2015).

[51] We exclude transactions between £125k and £145k so that our estimates are not biased by the removal of the £125k price notch in December 2014. As the notch has led to a lower transaction density just over £125k its removal might have led to increased transaction volumes in the price range just above.

between actual and counterfactual time trends (near April 2015) with the timing response to the LBTT introduction. Our results are presented in Figure 2.11 and Figure 2.12.

In the £145k to £333k price range, there seems to be some evidence that buyers delayed transactions in response to the anticipated introduction of LBTT (see Figure 2.11). In February and March, transaction volumes are (significantly) below counterfactual ones, however the actual level of transactions is not low in comparison to previous winters. The effect therefore does not seem to be very large, a decrease in transactions of roughly 16-17% in the months February and March. Nevertheless, when compared to the size of the tax saving from delaying (average tax savings are roughly £397, or 0.17% of average property value in this price range), this response is considerable, a 42.7% reduction in transaction activity per £1,000 reduction in tax liability.

In the over £333k price range, the impact of the behavioural response of bringing transactions forward is evident in the comparison of the time trends (see Figure 2.12). Compared to the counterfactual, March 2015 has approximately three times as many transactions, indicating a large change in transaction volumes in response to the introduction of LBTT. In this market segment, the increase in transaction volumes per £1,000 saved is around 26%[52]. Conversely, transaction levels in April and May are quite low (buyers who transacted in March do not transact in these months) although the market does seem to recover by June. Indeed, the missing mass of property transactions after April seems to be smaller when compared to the excess bunching before the tax regime change. For a surplus of 1,104 transactions in March, we only have 423 transactions missing in the next three months. There are two ways we can rationalise this finding. First, the additional transactions (or at least a share of these) might be a consequence of extensive margin

[52] This calculation is based on the average tax saving of £9,010 in this price range. Calculations of tax savings are based on tax savings for the actual transactions that took place during the announcement period (and the two months after). Once taking into account the average tax savings from re-timing, the timing response in the lower end of the market seems to be larger than that in the higher end. Note however, that delaying transactions should be a lot less costly than bringing them forward, the two responses are therefore not directly comparable.

responses, i.e. property transactions that would not have taken place if not for the tax incentives created by the time notch. Second, it is possible that the tax incentives created by the time notch were so salient that agents brought forward transactions from months way beyond April 1. Being incorporated into the counterfactual, these missing transactions could downward bias counterfactual estimates of transaction volumes for the final part of 2015. A solution to this issue would be to extend the excluded region to months near the end of 2015, however doing so would risk bias from the announcement and introduction of ADS in 2016 (see Section 2.2).

Overall, our evidence suggests that the Scottish property market responds quickly and drastically to tax incentives created by anticipated changes in transaction tax rates. From a tax revenue perspective, the presence (and extent) of timing responses leads to short-term losses in tax revenues as transactions get taxed under the more generous tax regime. Indeed, calculating tax revenues under actual and counterfactual scenarios for both price ranges gives us an estimated £18.38 million of lost tax revenues due to timing responses.

Figure 2.11: *Actual and counterfactual time trends of transaction volumes between £145k and £333k*

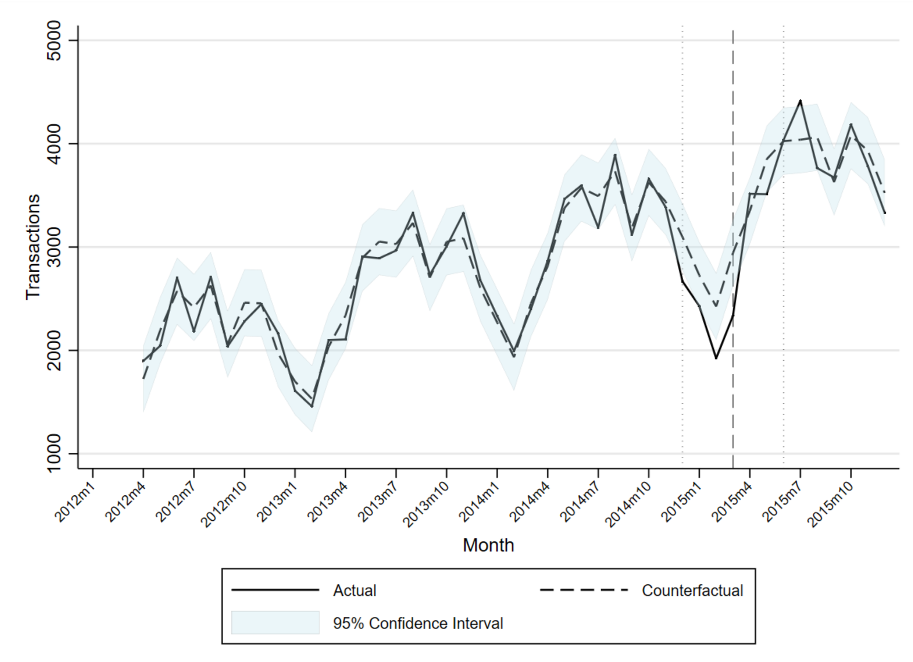
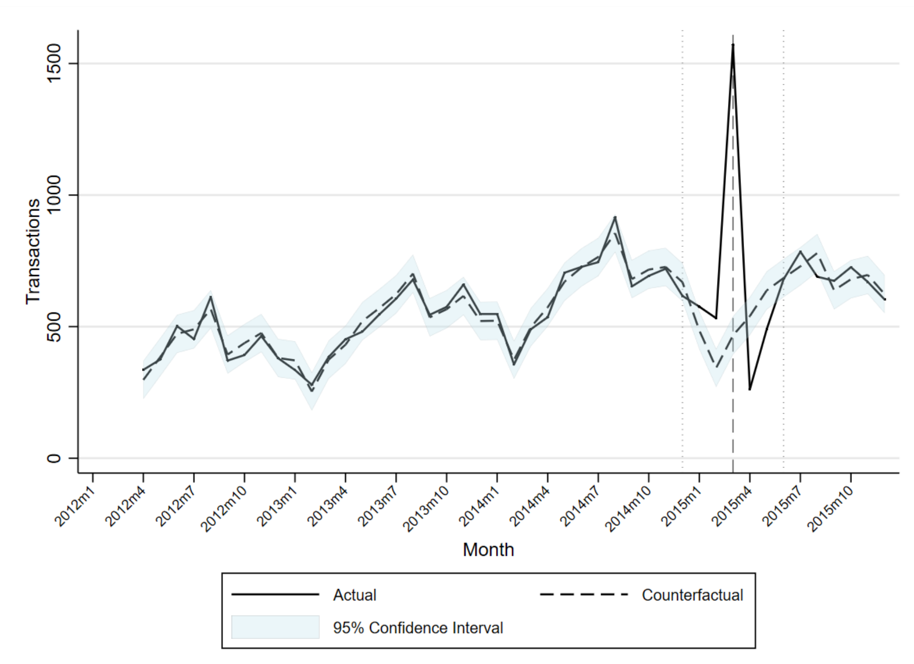


Figure 2.12: *Actual and counterfactual time trends of transaction volumes over £333k*



Notes: Solid lines represent the actual number of transactions in each month. Dashed lines represent the counterfactual number of transactions, calculated using predicted values from the regression model outlined above. The blue shaded area around the counterfactual represents the 95% confidence interval of the counterfactual estimates. The vertical dotted lines mark the beginning and end of the excluded region. The vertical dashed line marks the introduction of LBTT (time notch).

2.5.3 Estimating the permanent effects of progressive tax reform on transaction volumes

In this section, we assess the impacts of LBTT on property market activity in Scotland to investigate the permanent effects of progressive transaction tax reform on different market segments. Permanent effects arise because lasting changes in transaction tax rates change the incentives of buyers and sellers to transact, thereby affecting how often properties are bought and sold. For example, higher tax rates might permanently 'lock-in' property transactions that would have occurred under more generous transaction tax regimes (or in the absence of transaction taxes).

Progressive reform in Scotland should introduce both positive and negative distortions to the property market. This is because different price ranges were affected differently by the introduction of LBTT: while some lower price ranges saw reductions in tax rates, in higher price ranges tax rates increased. The main aim of the reform was to encourage home ownership among buyers in the lower price ranges and redistribute the tax burden to those transacting at the higher end of the market, more in line with their ability to pay[53]. Optimally, to minimise market distortions and maximise tax revenues, the positive response from the lower end of the market should be more substantial relative to the negative response in the higher end of the market. To assess whether this was the case, we investigate the impacts of progressive reform through a comparison, over time, of the price ranges affected and unaffected by the introduction of LBTT. Our baseline model is as follows:

[53] According to the Scottish Government, LBTT rates were unchanged or lower in comparison to previous stamp duty rates for roughly 90% of home buyers (Scottish Government, 2015).

$$\ln T_{im} = \alpha + \beta_1 Post_m + \beta_2 Post_m * LBTT_{CUT} + \beta_3 Post_m * LBTT_{INCREASE} + \sum_{i=1}^{4} \gamma_i Pre_m * Bunching_{BELOW} + \sum_{i=1}^{4} \delta_i Pre_m * Bunching_{OVER} + \beta_4 Bunching_i + \sum_{m=1}^{3} \gamma_m Announcement_m * Timing_{LOW} + \sum_{m=1}^{3} \delta_m Announcement_m * Timing_{HIGH} + \theta_y + \theta_m + \theta_i + \varepsilon_{im} \tag{2.3}$$

where our outcome variable $\ln T_{im}$ is the log number of monthly transactions in each £5,000 price bin 'i'. The dummy $Post_m$ is an indicator for the post-LBTT (treatment) period, while the interaction terms $Post_m\ x\ LBTT_{CUT}$ and $Post_m\ x\ LBTT_{INCREASE}$ indicate exposure to the LBTT reform in the relevant price ranges. The coefficients β_2 and β_3 are our coefficients of interest: β_2 measures the impact of tax cuts from the tax reform in the lower price ranges; whilst β_3 measures the impact of tax increases in the higher ranges. We also include several dummy variables and interaction terms to control for:

- the effects of bunching near notches in the previous SDLT period by interacting the SDLT period with the bunching price regions using the terms $\sum_{i=1}^{4} \gamma_i Pre_m * Bunching_{BELOW}$ and $\sum_{i=1}^{4} \delta_i Pre_m * Bunching_{OVER}$;
- the impact of bunching near notches and kinks in general through the dummy term $\beta_4 Bunching_i$ which takes a value of one for prices in potential bunching regions (see Section 2.5.1) ;
- and the impact of the LBTT announcement for the months affected by timing responses by interacting the announcement period with the price regions where timing responses were incentivised (see Section 2.5.2). We control for the timing effects through the interaction terms $\sum_{m=1}^{3} \gamma_m Announcement_m * Timing_{LOW}$ and $\sum_{m=1}^{3} \delta_m Announcement_m * Timing_{HIGH}$

Year and month fixed effects are also included to control for year and month-level trends in the outcome variable; while price bin fixed effects are used to control for unobservable differences in transaction market activity along the price distribution. Standard errors are clustered at the price bin level to allow for the correlation of error terms within price bins over time.

Our identification strategy relies on a panel difference-in-differences (diff-in-diff) analysis in which we compare changes in transaction numbers across price ranges affected (treated) and unaffected (control) by the LBTT reform. Selection of treated and control price ranges is based on the comparison of tax liabilities for equally priced properties, under the LBTT and previous SDLT regimes. Here, we consider the reform to be a change from the previous SDLT system to the new, more progressive LBTT regime (see Section 2.5.2). Figure 2.13 below summarises the differences in tax liabilities under the current LBTT and previous SDLT tax regimes. As shown in Figure 2.13, we can split up our sample into three distinct price segments: the £40 to £125k range where tax liabilities (for equally priced properties) were unaffected by the policy change; the £125k to £380 range where tax liabilities decreased under LBTT; and the £380k and above range where tax liabilities increased. In our analysis, we use the price range £40k to £125k, unaffected by the policy change, as the 'control' group. We exclude the price range below £40k to avoid any bias from the 2016 introduction of the Additional Dwelling Supplement (ADS), an additional surcharge payable for all second homes priced above £40k[54]. Our 'treated' price ranges, those affected by LBTT reform, are the price range £125k to £380k, where tax liabilities decreased; and the £380k and over price range, where tax liabilities increased. Average tax savings from LBTT in the 'tax cut' range were equal to roughly £1,600 (or 0.8% of average property value in this market segment), while average losses from LBTT in the 'tax increase' range were equal to roughly £7,800 (or 1.2% of

[54] The introduction of ADS should affect all price ranges to an equal extent over the price limit of £40k. This is also evident from the timing responses to the ADS introduction, which were substantial in all three distinct price ranges (see Figure 2.14 below). Nonetheless, to check whether our results are sensitive to the inclusion of the ADS period in our sample, we run an additional robustness check below.

property value)[55]. However, also note that the tax cuts affected a much higher number of potential transactions when compared to the tax increases[56].

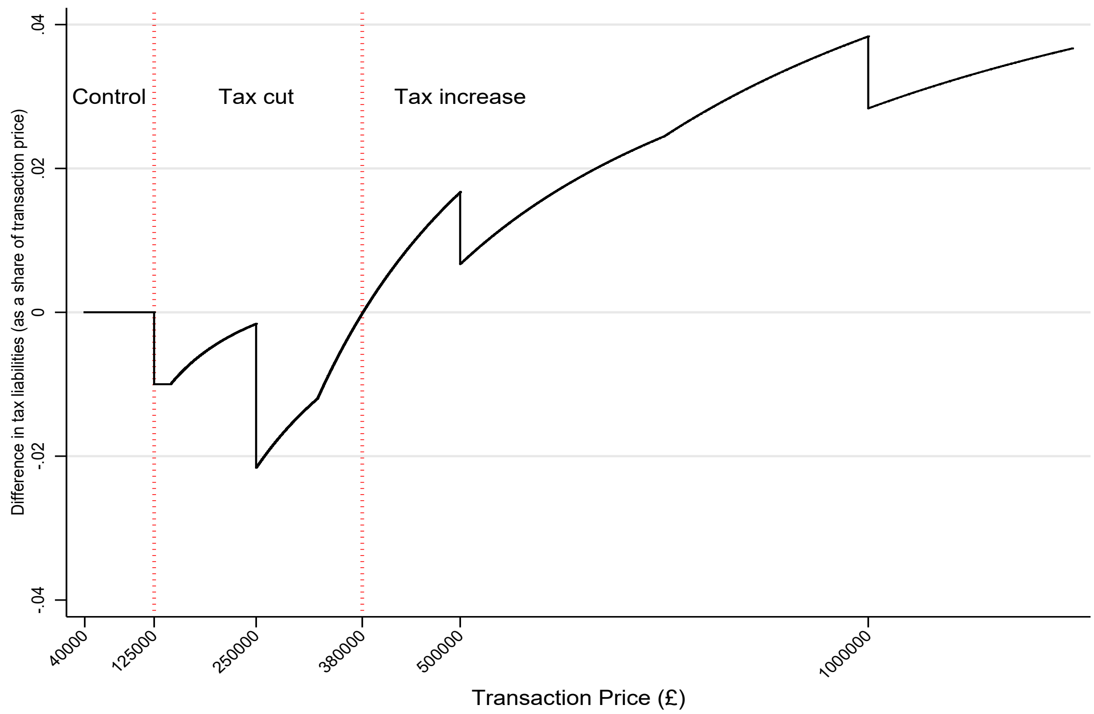
In Figure 2.14 and Figure 2.15, we plot trends in transaction volumes over time to evaluate the parallel trends assumption, which requires time trends in log transaction numbers to follow a similar pre-reform path for treated and control ranges. Figure 2.14 plots the log number of transactions for each of our price ranges, while Figure 2.15 plots the normalised log number, which is calculated by subtracting the pre-reform mean log number of transactions from each series (see Best and Kleven, 2018). The latter measure is used to facilitate the comparison of the time series by adjusting to a common scale. Both figures indicate nearly perfectly parallel trends before the reform between the control and treated price ranges, except from a small divergence in mid-2014 (see Figure 2.15)[57]. Timing responses to the introductions of LBTT and ADS are also observable in the large spikes in transaction activity in March 2015, and March 2016, respectively.

[55] These calculations are based on the estimated tax savings for the actual transactions that occurred during the LBTT period.

[56] According to the estimates of the Scottish Government, 90% of home buyers pay lower tax (or the same tax) under LBTT (Scottish Government, 2015). According to our own calculations, in the post-LBTT period, roughly 52% of transactions paid lower taxes than they would have under the previous system, 5% paid higher taxes, and the remaining 43% paid the same amount (usually zero).

[57] We further test for the validity of the parallel trends assumption through event study estimations below.

Figure 2.13: *Differences in tax liabilities along the price distribution - a comparison of the previous SDLT and current LBTT regimes*



Notes: The dotted vertical lines indicate the break-even prices of £125k and £380k: below £380k (and over £125k) liabilities are lower under LBTT in comparison to the previous SDLT system. Over £380k tax liabilities are higher under LBTT. The price range between £40k and £125k is unaffected.

Figure 2.14: *Time trends in transaction volumes - by price range*

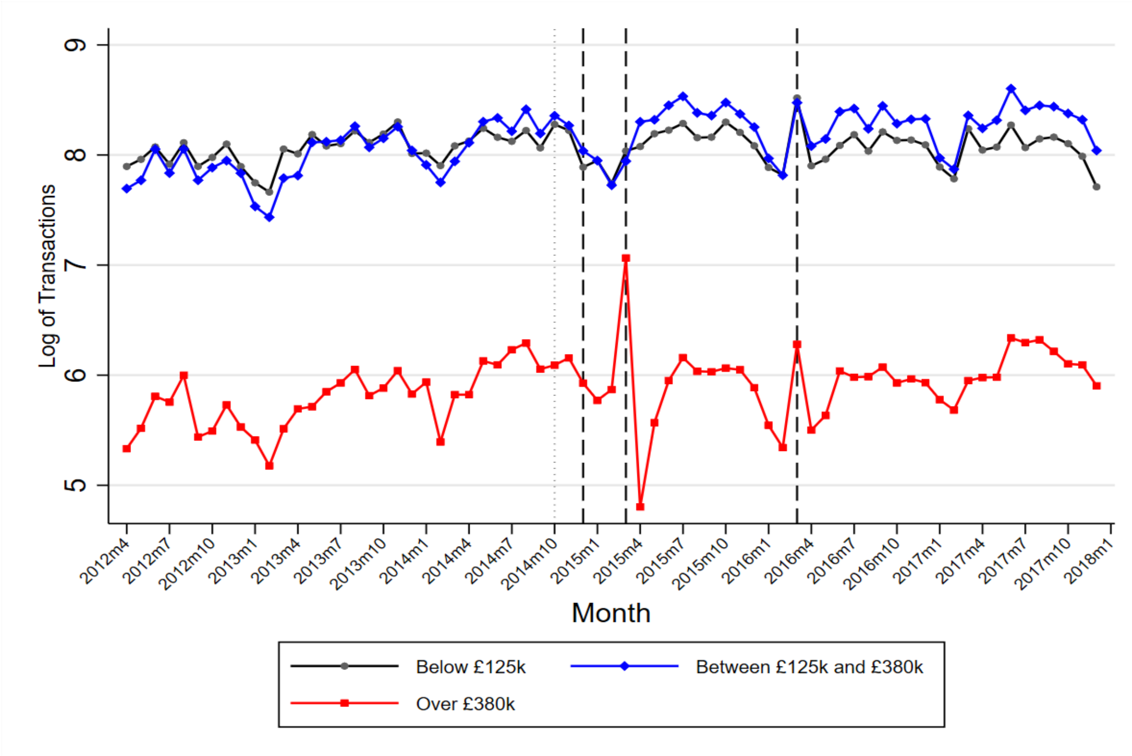
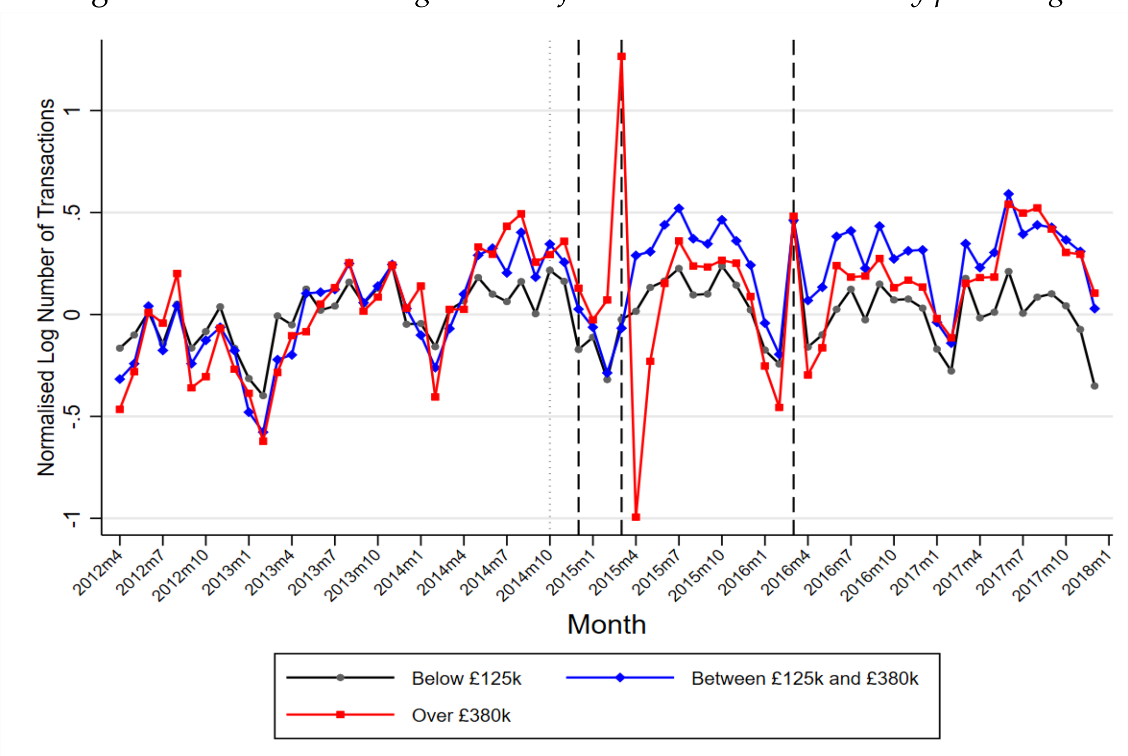


Figure 2.15: *Normalised log number of transactions over time – by price range*



Notes: Normalised log transaction volumes are calculated by subtracting the pre-announcement mean from each price range series. The vertical dotted line before October 2014 indicates the initial LBTT announcement; the vertical dashed lines indicate the period between the announcement of LBTT rates and their introduction; and the vertical dashed line before April 2016 indicates the introduction of ADS.

Our diff-in-diff model is estimated using OLS and results are presented in Table 2.4. Several robustness checks are carried out to test the sensitivity of estimates to changes in the model specification. First, we exclude the £10k price ranges around the £125k and £380k tax liability thresholds to control for possible spillovers between neighbouring price ranges[58]. This robustness check is also necessary to address concerns about potential endogeneity, i.e. home buyers deliberately transacting in certain price regions to qualify for tax cuts. Second, we exclude months from after December 2015 to check whether our estimates are sensitive to the introduction of ADS. Finally, we disaggregate our data at the Scottish Postcode Area level and use area fixed effects to control for potential regional heterogeneity[59]. Results from robustness checks are presented in Columns 3-8 of Table 2.4.

Overall, the estimates presented in Table 2.4 are indicative of a significant (and robust) aggregate level effect of the reform over the price range £125k to £380k, where transaction taxes were reduced. The point estimates for this price range span between 0.174 and 0.288, indicating that on average and relative to the control group, the LBTT reform has led to a permanent increase in transaction volumes of 17.4 to 28.8%, or equivalently, an increase of 10.9 to 18% per £1,000 reduction in tax liability. These findings provide robust evidence that progressive reform did encourage transaction activity in the market segment where it was intended to[60]. Conversely, estimates relating to the £380k and over price range, where tax liabilities increased under LBTT, are small and mostly insignificant. The only significant point estimate is for the baseline model (Column 1), showing a positive impact from LBTT reform, but this finding lacks robustness since point estimates

[58] One potential source of spillover effects may be the presence of real estate chains, i.e. buyers of properties simultaneously selling existing properties at slightly different prices leading to an increase in transaction activity in adjacent price ranges (Best and Kleven, 2018).

[59] Disaggregating by postcode areas reduces the number of transactions we can include by approximately 3% due to missing data on postcode areas for some of our sample. Missing information on postcode areas does not seem to be correlated with where a property is located, suggesting that there is no structural bias when disaggregating.

[60] Indeed, even the lower bound of our lowest point estimate suggests a 9.4% increase in transaction volumes in response to the tax cut.

either change sign or are insignificant in other specifications.

2.5.3.1 Event study specification

In this section, we modify our baseline specification in Equation 2.3 and interact our reform exposure dummies with month fixed effects. This specification, a variant of an event study design, serves two purposes: 1) interacting the reform exposure dummies with all post-reform months allows us to assess the reform's effect over time and 2) pre-reform point estimates allow us to evaluate whether there are significant pre-treatment differences in transaction activity between treated and control price ranges.

In this specification, point estimates for the treated price ranges correspond to the relative deviations from the control group, conditional on covariates and fixed effects, for each month of the sample. These point estimates are plotted in Figure 2.16 and Figure 2.17, for the 'tax cut' and 'tax increase' price ranges, respectively. We exclude the price range between £333k and £380k as this might have been differentially affected by timing and permanent responses[61].

In Figure 2.16, the positive impact of the reform on transaction volumes is evident from the consistent (and significant) upward trend in point estimates after the introduction of LBTT. Significance is indicated by the 95% confidence intervals (vertical bars) not spanning zero, and is achieved for all but three of the post-reform months. The impact of LBTT also seems to be a long-term one, given that it does not diminish over time. In the pre-announcement period, there are only two significant point estimates. This is not a threat to the parallel trends assumption since estimates for all other months are insignif-

[61] This is because in this price range, there was a short-term incentive to bring transactions forward before April 2015 but the permanent change (when comparing the previous stamp duty system and LBTT) was a tax cut. Strangely, this means that for home buyers ready to transact in the months near April 2015 there was a short-term incentive to 'avoid' the permanent tax cut and pay even lower taxes on their transactions under the temporary tax regime.

Table 2.4: *OLS Results – Difference - in - Differences Estimates*

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	Baseline	*Baseline*	*£10k excl.*	*£10k excl.*	*Until 2016*	*Until 2016*	*Area FE*	*Area FE*
$\beta_2 Post_m * LBTT_{CUT}$	0.288***	0.274***	0.270***	0.263***	0.282***	0.271***	0.206***	0.174***
	(0.025)	(0.056)	(0.031)	(0.057)	(0.030)	(0.054)	(0.022)	(0.040)
$\beta_3 Post_m * LBTT_{INCREASE}$	0.077***	0.041	0.049	0.019	0.018	-0.007	0.002	-0.044
	(0.025)	(0.042)	(0.031)	(0.043)	(0.036)	(0.048)	(0.024)	(0.039)
Observations	11,481	11,481	11,276	11,276	7,531	7,531	72,761	72,761
R-squared	0.955	0.951	0.956	0.952	0.954	0.950	0.745	0.741
Controls	Yes	No	Yes	No	Yes	No	Yes	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price bin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	No	No	Yes	Yes

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the price bin level. Columns 1 and 2 show estimates for our baseline specification in model 1. Columns 3 and 4 show estimates for the robustness check where £10k price range around kink points are excluded to control for spillover effects between price ranges. Columns 5 and 6 show estimates for the robustness check where months after 2015 are excluded. Columns 7 and 8 show estimates for the robustness check where we include postcode area fixed effects to account for regional heterogeneity in transaction market activity. The level of observation in Columns 1 to 6 is price bin - month. In Columns 7-8 the level of observation is postcode area - price bin - month.

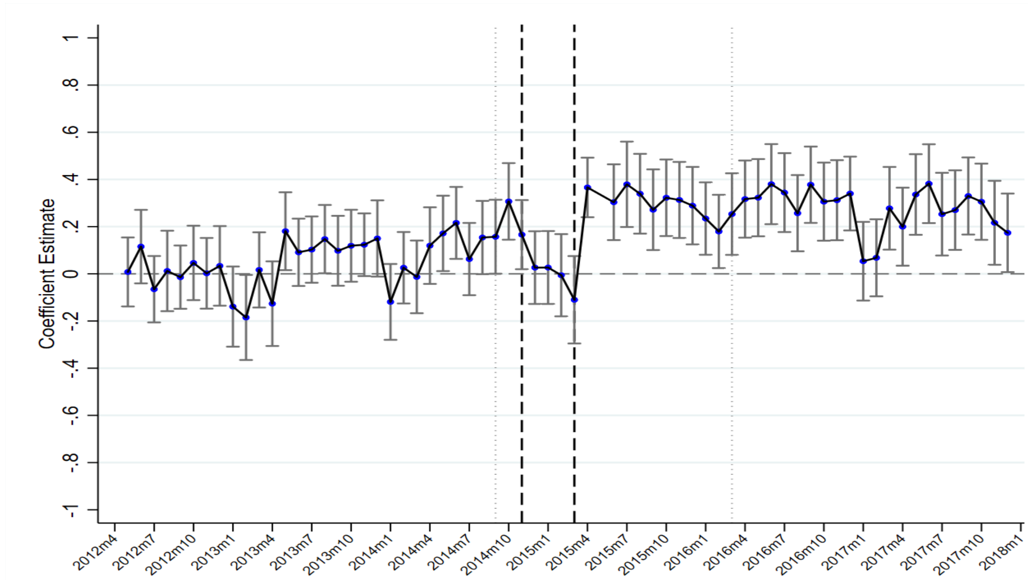
icant and the significant estimates for outlier months might be random or a consequence of some exogenous shock to transaction activity in the treated price ranges[62].

For the over £380k price range, the event study estimates provide no evidence of a significant post-reform impact on transaction activity (see Figure 2.17). The only significant point estimates are for months immediately after the introduction of the new tax regime, in a period where markets are likely still recovering from the timing responses in the previous months (see Section 2.5.2)[63]. Pre-announcement point estimates are insignificant and close to zero, indicating no violation to the parallel trends assumption. Overall, while most point estimates in the post-reform period are negative, indicating reduced transaction volumes from progressive reform, neither of these estimates are significant and therefore we find no evidence of a permanent effect on the price range where tax rates increased.

[62] A potential source of bias in this period is from the introduction of the Help-to-Buy (HTB) scheme, a government initiative designed to provide small loans for properties under £400k. The scheme was started in October 2013 and should impact all price ranges below the limit to an equal extent (while higher ranges were excluded as wealthy home buyers have more ability to pay). To check if HTB biases our results, we run a robustness check where we include an HTB dummy for the affected price ranges and time period. This modification does not change our results to any extent.

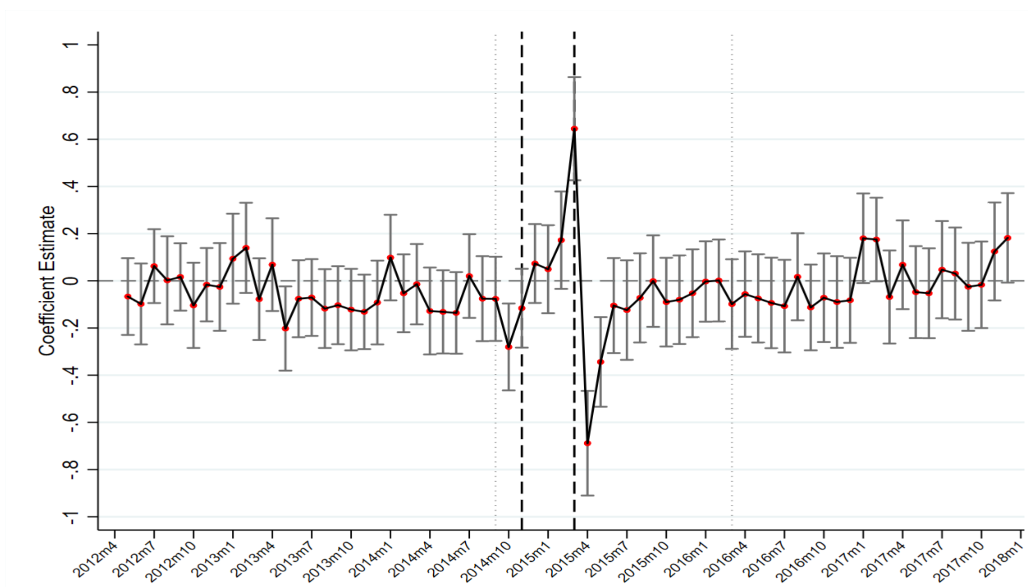
[63] Both the timing response in March 2015, and the subsequent missing transactions are evident from the event study plot in Figure 2.17. Note, however that the point estimates are not directly comparable to our estimates in Section 2.5.2, as the latter use a different sample and counterfactual.

Figure 2.16: *Event Study Estimates - £125k to £333k price range*



Notes: This plot shows point estimates for each month interacted with the treated price range £125k to £333k. Significance is indicated by the 95% confidence interval (vertical bars) not spanning zero. The vertical dotted line before October 2014 indicates the initial LBTT announcement; the vertical dashed lines indicate the period between the announcement of LBTT rates and their introduction; and the vertical dotted line before April 2016 indicates the introduction of ADS.

Figure 2.17: *Event Study Estimates - £380k and over price range*



Notes: This plot shows point estimates for each month interacted with the treated price range £380k and over. Significance is indicated by the 95% confidence interval (vertical bars) not spanning zero. The vertical dotted line before October 2014 indicates the initial LBTT announcement; the vertical dashed lines indicate the period between the announcement of LBTT rates and their introduction; and the vertical dotted line before April 2016 indicates the introduction of ADS.

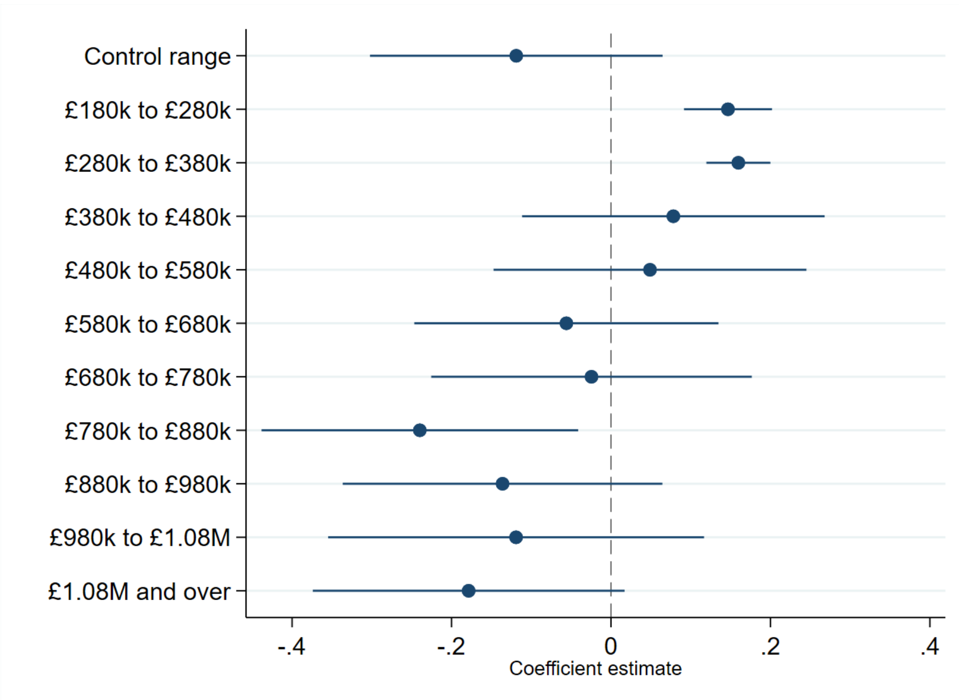
2.5.3.2 Testing for the heterogeneity of treatment effects across price ranges

Our third specification attempts to quantify the property market effects of LBTT reform along the price distribution. To identify these effects, we modify our baseline specification in Equation 2.3 by interacting the post-reform dummy $Post_m$ with price range dummies for £100k price ranges on both sides of the £380k tax savings cut-off (see Figure 2.13).

Results are summarised using the coefficient plot in Figure 2.18. The plot shows point estimates for each price range interacted with the post-reform dummy, along with 95% confidence intervals. Confidence intervals (horizontal lines) not spanning zero are indicative of significant estimates. In our estimation, we also include a placebo test for the control price range. Naturally, the reform should not have a significant effect on the unaffected parts of the price distribution. Indeed, Figure 2.18 provides evidence that this was not the case, as the reform's impact on the control price range is not significant at the 5% level.

Overall, the results summarised in Figure 2.18 tell a clear story: point estimates get smaller, and change sign as we move through the price distribution. The initial positive (and significant) effects from the reform in the lower price ranges (where tax rates decreased) are not surprising given that our aggregate level estimates from the last two sections were indicative of a similar positive impact. In a similar vein, the observed negative point estimates for the higher end of the market, albeit imprecisely estimated apart from the £780k to £880k range, are perfectly sensible given that progressive reform led to the largest increases in tax rates in this market segment. However, this negative impact was only really felt in the top end of the market, as our estimates suggest that the price range £380k to £780k was largely unaffected. Since the top end of the market (over £780k) contributes a mere 0.64% of total transactions in Scotland over our sample period, aggregate level negative distortions from the reforms were likely rather small.

Figure 2.18: *Treatment Effect by Price Range*



Notes: This plot shows point estimates for each price range interacted by the LBTT (post-reform) period. Significance is indicated by the 95% confidence intervals (horizontal lines) not spanning zero. Statistically significant differences between point estimates are indicated by the respective confidence intervals not spanning each other. The range £125k to £180k is excluded to avoid multicollinearity.

2.6 Discussion

In general, taxation of property transactions is not optimal (Mirrlees and Adam, 2010). In Scotland this sub-optimality currently arises for (at least) two reasons: 1) transaction taxes discourage an otherwise beneficial economic activity and 2) agents are highly responsive to them, leading to large welfare costs (deadweight losses) from behavioural responses. Nonetheless, our results suggest that recent policy changes in Scotland were successful in limiting overall distortions from the transaction tax system. The following paragraphs reflect on our findings in more detail.

a) *Price notches and kinks*

Our analysis of the previous stamp duty regime shows that the 'notched' design of that system had a considerable distortionary impact on the Scottish property market showing up in bunching of transactions. These bunching responses are likely a result of agents manipulating prices in order to move into lower tax bands and avoid the abrupt increase in tax liabilities at specific price notches. Overall, our findings are in line with the predictions of the theoretical literature and prior empirical results (Kopczuk and Munroe, 2015; Kleven, 2016; Slemrod et al., 2017; Best and Kleven, 2018).

Contrary to the case of the previous SDLT regime, we find only limited evidence of bunching under the 'kinked' LBTT system. Bunching responses are only apparent in transactions close to the £250k and £750k kinks, and magnitudes are smaller than for the bunching responses observed at notches. Possibly, the fact that we observe bunching responses only at these kink points and not others (£145k and £325k) is a result of these kinks providing more salient reference points for agents to strive towards (see Kleven, 2016). Overall, we can say that the removal of price notches led to a simple and more sensible tax system, where manipulation of prices to avoid paying higher taxes is incentivised to a far lesser extent. Consequently, our evidence suggests that replacing the notched structure of stamp duty with a kinked one was a sensible policy of the Scottish Government, and given the responsiveness of agents and the associated distortions to the property market, our recommendation is that notches should be avoided in future stamp duty regimes.

b) *Time notches*

Our findings in Section 2.5.2 provide evidence that Scottish agents are highly responsive to the presence of time notches and will time transactions in order to realise tax savings. These responses result in lower tax revenues for the government as the inter-temporal substitution of transactions leads to more transactions being taxed under the more generous tax regime. In Scotland, time notches are likely to continue to emerge upon the introduction and announcement of new tax policies. The savings opportunities provided by these notches are usually present for roughly four months as policies announced in December are generally implemented the following April – announcement windows of this length provide ample time for agents to re-time transactions. The current devolved context means that Scottish budget setting procedures are complicated by their reliance on UK budgetary procedures (Eiser, 2017) – institutional constraints prohibit the government from creating shorter announcement windows for policy changes. Consequently, in the Scottish case, distortions from time notches are an unintended consequence of the devolved fiscal setting. As long as these institutional arrangements are in place, transaction tax changes are going to be anticipated, announcement periods will remain lengthy, and timing responses will likely continue to cause large temporary distortions to market activity. To mitigate the effects of time notches on the property market, policy makers need to account for the nature and extent of associated distortions when designing and implementing new policies. Our estimates of timing responses (and corresponding estimates for behavioural elasticities) could therefore be useful to policy makers attempting to forecast the distortionary effects of future policy announcements.

c) *Permanent effects of progressive transaction tax reform*

Our finding that progressive LBTT reform has led to an increase in transaction activity in the lower end of the market is consistent with the government's objective to encourage transaction activity in lower price ranges. The underlying economic rationale behind this objective is to expand homeownership and facilitate access to the housing market to first-time buyers and younger households (see Scottish Government, 2015). One explanation for the sizeable effect observed in this price range is that, according to the theoretical model in Best and Kleven (2018), low-income home buyers (especially first-time buyers) are likely to be highly leveraged and constrained by mortgage downpayments, increasing their responsiveness to transaction taxes that need to be paid upfront. It is therefore possible that higher transaction tax rates in the past led to a permanent 'lock-in' effect, with transactions in lower price ranges not taking place due to the pressure transaction taxes put on downpayment constrained agents with low incomes.

Our findings also suggest that progressive tax reform did not have an overall significant negative effect on the transactions in the higher price ranges where tax rates increased. More substantial negative effects from the tax reform, albeit in most cases imprecisely estimated, only appear towards the most expensive segment of the market (over £780k) where less than 1% of total transaction activity takes place.

From a policy perspective, these findings are promising, as the boost to market activity at the lower end of the price distribution was not offset by a similarly sized negative response at the higher end of the market. In other words, the intended positive distortions introduced by progressive reform are seemingly larger, in relative terms, than the negative ones. Assuming that institutional arrangements render abolition (or comprehensive cuts) infeasible, this finding suggests that progressive reform could potentially be a 'second-best' policy option that can achieve efficiency gains for the vast majority of property market participants without causing unwanted distortions in transaction volumes and tax revenues in the other parts of the market[64]. In general, whether there are efficiency gains to be made from progressive reform should depend on the relative responsiveness of different market segments to changes in tax rates – this is equivalent to an 'inverse-elasticity rule' which would postulate that, to minimise distortions, more responsive market segments should be taxed at lower rates and *vice versa* (see Sandmo, 1976).

Nonetheless, some important caveats are worth noting before we use these results as the basis for generic policy recommendations. First, it is unclear whether the differences in responsiveness we observe across the price distribution are a consequence of transactions in lower price segments being more responsive to all types of tax changes than higher ones, or the lower segment being more responsive to a tax cut than the higher one is to a tax increase. If agents across the price distribution are more responsive to tax cuts than they are to tax increases, then responsiveness is endogenous to the sign of the tax change and the inverse-elasticity rule becomes self-fulfilling. Furthermore, even if agents were equally responsive to tax cuts and tax increases, it would be difficult to use our results to determine the optimal degree of progressivity for the transaction tax system. What our analysis of Scottish reform tells us is that in aggregate, the responses to progressive reform are consistent with a desirable policy outcome. It is also plausible that further increases in

[64] One potential caveat to this argument is that the tax revenue implications of the negative effect observed on the range of very expensive properties may still be substantial as these properties account for a disproportionate share of the tax revenues from LBTT. For example, in 2017-18, the over £750k price range accounted for nearly 16% of overall tax revenues from LBTT (Revenue Scotland, 2017).

progressivity would have a similar effect. In fact, any increase in progressivity will likely lead to an efficiency gain given the differences in marginal responsiveness to changes in progressivity along the price distribution. As long as there are differences in marginal responsiveness, tax cuts in some part of the price distribution will be more (positively) distortive to transactions than tax increases in another part. In the case of Scotland, we can therefore conclude that while initial progressive reform has led to some efficiency gains via the large relative stimulus effect on the lower end of the market, it is unclear whether further improvements could be made and more advanced theoretical and empirical frameworks are needed to assess this.

Overall, our evidence on permanent effects suggests that due to the responsiveness of Scottish agents, cuts in transaction taxes can be used rather effectively to stimulate the property market, even outside periods of economic downturns[65]. Furthermore, our evidence on the effects of progressive reform suggest that progressive changes in transaction tax schedules could also be effective in achieving redistributive goals given 1) minimal distortions to the market segments where tax rates increase and 2) the boost to transaction activity in the lower end of the market, which includes the majority of Scottish home buyers (see Section 2.5.3). These results may be driven by a number of desirable economic mechanisms. For example, reduced moving costs due to lower transaction tax rates could facilitate job search by making it less costly to relocate in search of employment (Mirrlees and Adam, 2010). Increased mobility may also mean that lower tax rates encourage retired households to downsize, leading to a better allocation of the housing stock in terms of the optimal amount of space needed for households (Glaeser and Luttmer, 2003). Nonetheless, our analysis does not directly identify these effects and further research is needed to examine the mechanisms that drive our results.

[65] For evidence on the fiscal stimulus effect of transaction tax cuts during periods of economic downturns, see Besley et al. (2014) and Best and Kleven (2018).

2.7 Conclusions

This chapter studied the distortionary impact of four distinct features of recent Scottish transaction tax systems: 1) the presence of price notches in the previous stamp duty system 2) the existence of kink points in the LBTT system 3) the appearance of a time notch corresponding to the April 2015 introduction of LBTT and 4) the shift to the more progressive LBTT regime.

Using a bunching estimator methodology, whereby we estimate counterfactual transaction densities for each price range, we find clear evidence of the bunching of transactions around price notches in the previous stamp duty system. In contrast, we find limited evidence of bunching at (some) kink points under LBTT, but these responses are smaller and less significant than the bunching found in the case of the previous (notched) transaction tax system. We further find substantial evidence that the temporary tax saving opportunity created by the early announcement of LBTT rates resulted in large-scale re-timing of property transactions. Finally, studying the effects of progressive reform across the entire price distribution, we find that LBTT had a substantial positive effect on transaction activity in the market segment in which tax rates decreased; and only a limited (and not overwhelming) negative effect at the higher end of the market where tax rates increased, indicating that the latter part of the market was less responsive to progressive tax reform.

Overall, our results suggest that the Scottish property market is highly responsive to changes in transaction taxes, and is particularly responsive to tax saving opportunities available due to the presence of notches in the tax system. While, overall, transaction taxes are distortionary to property market behaviour, tax policy should focus on mitigating their effects on economic behaviour. Removing price notches from the Scottish tax system was therefore a sensible policy, but time notches continue to emerge due to institutional arrangements related to Scotland's devolved policy setting. In addition, behavioural responses to the LBTT reform suggest that progressive changes to transaction tax system should be encouraged if they induce transaction activity in lower market segments more than they reduce activity in higher segments.

The evidence from our analysis also suggests that due to the high degree of responsiveness to transaction tax changes, cuts to these taxes can be used very effectively to stimulate activity in the property market. This boost to the property market is likely driven by behavioural responses that are desirable for the housing and labour market outcomes of households. Future research should focus on directly identifying these household level behavioural responses to changing transaction tax rates.

Chapter 3

THE IMPACT OF HOUSING SUBSIDY CUTS ON THE LABOUR MARKET OUTCOMES OF CLAIMANTS: EVIDENCE FROM ENGLAND

3.1 Introduction

Housing subsidies assist low-income households with their rent obligations to help them afford appropriate housing. Critics note that these subsidies might depress work incentives among claimants leading to a reduction in labour supply. Following this line of reasoning, cuts to housing – and other in-kind – subsidies are often justified by governments as measures to encourage work or search effort by claimants (Taylor-Gooby, 2012). In England, recent housing subsidy cuts were justified in a similar manner, the main government objective being the creation of a 'fair' system where claimants are encouraged to work in order to afford quality housing[66].

Despite the common justification, economic theory yields largely ambiguous predictions on the sign (and size) of labour supply impacts from housing subsidies (Murray, 1980; Schone, 1992; Moffitt, 2002; Shroder, 2002). The standard neoclassical model of labour supply predicts that housing subsidy provision will lead to a reduction in labour supply through both substitution and income effects (Moffitt, 2002). Conversely, alternative models predict that labour supply effects may vary depending on whether housing is a substitute or a complement to leisure (Murray, 1980); whether housing is a complement to other consumption goods (Schone, 1992); or could even be positive if subsidies lead to reduced housing uncertainty and allow claimants to spend more time seeking employment (Collinson et al., 2015).

While theoretical predictions on the labour supply impacts of housing subsidies are ambiguous, empirical evidence mostly confirms the stipulations of the neoclassical model. Recent quasi-experimental studies from the U.S. find conclusive evidence that housing subsidy provision has a negative effect on labour supply, although this effect diminishes with time and is mostly rather small (Mills et al., 2006; Jacob and Ludwig, 2012; Carlson

[66] See House of Commons, Work and Pensions Committee (2010) for an overview of the policy consultation.

et al., 2012). In contrast, empirical evidence from outside the U.S. is inconclusive and rather sparse (Shroder, 2010). Furthermore, the case studies in the literature tend to focus on the effects of housing subsidy provision, and very few studies focus on the effects of a reduction or withdrawal of subsidy entitlements from existing claimants. More specifically, when housing subsidies are withdrawn or reduced, does this induce claimants to increase labour supply either along the extensive (having a job) or intensive (hours of work) margin? In this chapter, we aim to answer this question through an analysis of recent housing subsidy cuts in England.

In England, through the Housing Benefit (HB) system, housing subsidies constitute a significant share of welfare expenditures. In the first decade of the 21^{st} century, expenditures on housing subsidies increased by 46%, with annual expenditures in 2010/11 totalling at £21.4 billion (Wilson et al., 2016). Along with a range of austerity measures aimed at reducing the public deficit, the 2010 Coalition Government introduced several changes to the English housing subsidy system[67]. Specifically, the June 2010 Budget announced changes to the way Local Housing Allowance (LHA) rates, which determine subsidy entitlements for claimants in the private rental sector (PRS), are set. These changes, rolled out in 2011-12, resulted in substantial housing subsidy cuts for PRS claimants. The government's justification for the LHA cuts relied on two underlying objectives. First, they intended to curb expenditures from the housing subsidy system. Second, they wanted to ensure that households on subsidies do not occupy more expensive housing than low-income working families. This second objective was an attempt to induce claimants to seek employment, and to create the means for housing through labour and not benefit income (see Wilson et al., 2016).

[67] The reforms were rolled out UK wide, however the changes discussed in this chapter concerned the English subsidy system to the greatest extent, as devolved administrations in other parts of the UK had some discretion over related housing policies and in some cases decided to mitigate the impacts of housing subsidy reform. For this reason, our analysis of the recent reforms to the housing subsidy system only concerns England.

The LHA cuts proved to be controversial right from their announcement. Critics of the new system commented on the hardships the subsidy cuts created for claimants by reducing rent affordability leading to increased risk of eviction by landlords. A briefing report released by the Shelter Foundation claimed that while the LHA cuts led to a deterioration in housing conditions for PRS claimants, in many cases leading to eviction and homelessness as claimants found themselves being unable to pay rent, they did not lead to a noticeable increase in employment (Shelter, 2015). There is also evidence from the academic literature that the subsidy cuts led to a substantial reduction in rent affordability, although evidence on whether the reforms reduced or increased claimant mobility is mixed (Brewer et al., 2014; Braakmann and McDonald, 2018) and no analysis to date has focused explicitly on the reforms' effect on labour market outcomes.

This chapter contributes to the empirical literature on the labour supply impacts of housing subsidies through a comprehensive analysis of the impacts of the recent LHA cuts on the labour market outcomes of private rental sector claimants in England. More specifically, using England as a case study, we aim to fill the gap in the literature and investigate the effects reduced housing subsidies have on the labour market outcomes of existing claimants. In terms of policy implications, our results shed light on the efficacy of housing subsidy cuts in encouraging labour market activity among claimants.

Our analysis draws on individual-level panel data on housing subsidy claimants in England from the Understanding Society (US) longitudinal survey for the period 2009-2015. Our empirical strategy makes use of a panel difference-in-differences (diff-in-diff) approach whereby we assess the effects of the LHA cuts through a comparison, over time, of groups affected (treated) and unaffected (control) by the policy. As our control group, we use a cohort of social rental sector (SRS) claimants. Subsidy claimants in the control group are renting accommodation from local authorities and housing associations (as opposed to private landlords) and are therefore unaffected by the housing subsidy cuts. We display our estimates using an event-study design, whereby the reform's impacts can be assessed at different points in time and long-run impacts can be estimated.

Our results provide no evidence of a significant (and lasting) impact from the subsidy cuts on labour market outcomes either along the extensive or the intensive margin. Our findings therefore provide no evidence that the government's objective to get claimants back to work (or even get them to seek work) was accomplished, at least not at the aggregate level.

The remainder of this chapter is organised as follows. Section 3.2 outlines the recent changes to the UK housing subsidy system. Section 3.3 provides an overview of the relevant literature. Section 3.4 describes the data. Section 3.5 outlines our identification strategy. Section 3.6 presents and discusses the results. Section 3.7 concludes.

3.2 Policy Background

In England, all individuals can apply for Housing Benefit (henceforth HB), a means-tested subsidy that provides assistance with the rental costs of housing, as long as claimants: 1) live in rental accommodation 2) are on a low-income or are claiming benefits and 3) possess savings lower than £16,000. The amount of HB received by each individual/household is determined as follows:

$HB = \min\{rent,\ HB_{max}\}\ if\ Y \leq Y^T$ or

$HB = \min\{rent,\ HB_{max}\}\ - 0.65\left(Y - Y^T\right)\ if\ Y > Y^T$

where HB_{max} is the maximum eligible housing subsidy amount, Y is household income, and Y^T is the threshold income for HB eligibility. Above the threshold income, housing subsidies are withdrawn at the taper rate of 65%, i.e. the subsidy amount is reduced by 65 pence for every £1 increase in income.

In the PRS, low-income households can access housing benefits through the Local Housing Allowance (henceforth LHA) system. The LHA system was introduced on April 7, 2008 and provides a set of rules that determine the amount of housing subsidy low-income households are eligible to claim (see Wilson et al., 2016). The extent to which LHA will cover rental costs is determined by the private market rent distribution in the specific geographical area (referred to as Broad Rental Market Area); the size of the household; and the earnings and income from other benefits claimed by household members. Upon their introduction, LHA rates were set to cover housing costs for properties with values below the local median house price (the cheapest half of local properties). Rates were then adjusted (monthly) to reflect inflation in rents.

Starting in 2011, the UK Government introduced (as part of the Welfare Reform Act 2012) several subsidy cuts and eligibility rule changes to the LHA system:

- LHA rates were set so that they only cover the bottom 30^{th} percentile of local rental properties instead of the bottom 50^{th} percentile;
- the Shared Accommodation Rate (a lower rate for claimants not living in shared accommodation) was extended to cover a wider age group;
- LHA rates were capped;
- the £15 per week excess, the amount claimants could keep when their rent was below the LHA rate, was removed;
- LHA rates no longer adjust to the inflation of rental costs – they are currently uprated in line with CPI inflation[68]

[68] This is important because according to analysis by Shelter (2015) rents have been rising more sharply than LHA rates, particularly in London.

The introduction of the reform package was staggered: it was announced through the June 2010 budget and rolled out starting from April 2011, however, due to transitional protection periods some claimants were not rolled in until late 2012 (see Brewer et al., 2014). Transitional protection also meant that the earliest enrolment date for existing claimants was January 2012, and only new claimants were rolled in in the months before.

In the social rental sector (SRS), where accommodation is rented from local authorities or housing associations, there were plans to introduce the same LHA rules as in the PRS, but this policy was first deferred and then scrapped altogether by the UK Government (see Wilson et al., 2016). Instead, a tax on 'spare' bedrooms was introduced to encourage downsizing among claimants (Gibbons et al., 2018). This policy, often referred to as the 'bedroom tax', was introduced in April 2013, and constituted a small monetary 'punishment' for households occupying properties with more rooms than they are entitled to based on policy rules. The 'bedroom tax' therefore targeted a different policy base in comparison to the LHA cuts: it only led to a subsidy cut for those in the SRS not adhering to the specific subsidy eligibility rules, whereas the LHA cuts applied to all subsidy claimants in the PRS.

The objectives of the government with the LHA cuts were related to two concerning aspects of the housing subsidy system: its cost, and its effects on the housing and labour markets (Tunstall et al., 2015). First, a concern was that spending on housing subsidies constituted a large share of total welfare expenditures in the pre-reform period and yet did not contribute to an investment in the housing stock. This was because the subsidies were paid to PRS landlords for existing property. Second, as housing subsidies help pay rent costs for those at low wages, they effectively subsidise low wage (labour) income.

Considering these concerns, on one hand, the government's objective was to reduce expenditures from the housing benefit system and make the system simpler. On the other hand, they wanted to encourage labour market participation amongst claimants by providing more incentive to withdraw from benefits and seek work. As the Department for Work and Pensions (DWP) argued:

> Providing some customers, mainly in London, with the ability to live in very high cost rented properties makes it extremely unlikely they would ever move completely off Housing Benefit because of the very high income levels required. Moving to more affordable accommodation could therefore encourage households to take up employment and move completely off benefits[69].

During the consultation period, the main government justification of the subsidy cuts was based on the assertion that housing subsidy claimants occupy more expensive housing than working individuals not in receipt of benefits. The Minister of Pensions put it this way:

> Low-income households rent at about 90% of what the Housing Benefit recipients are renting at. So they are renting at a lower level. [...] The facts are that low-income people who are not taking Housing Benefit are having to live in cheaper housing[70].

The government argued that housing subsidy cuts were necessary to eliminate an 'unlevel' playing field, so that subsidy recipients would not enjoy higher quality housing than low-income working families, and would be encouraged to seek work. Some experts found this justification unsatisfactory. For example, a study by the Cambridge Centre for Housing and Planning Research (Fenton, 2010) claimed that the LHA cuts were likely to have no impact on the labour market activity of claimants as most subsidy recipients who were able to work were already in employment before the reforms. The author also

[69] See House of Commons, Work and Pensions Committee (2010).
[70] See House of Commons, Work and Pensions Committee (2010).

claimed that the reason unemployment levels were high in some areas was due to structural weaknesses in regional economies, and not due to individuals' lack of willingness to work. Qualitative assessment of the recent reform by Shelter (2015) also found that it had no noticeable impact on the labour market activity of claimants. Later sections of this chapter will aim to identify these labour market impacts empirically.

3.3 Related Literature

3.3.1 Justification for housing subsidy programs

The main motivation for housing subsidy programs is to provide affordable housing to low-income households (Collinson et al., 2015). By helping low-income households afford appropriate accommodation, housing subsidies also leave more income for the consumption of other necessary goods that might have previously been under-consumed. Housing subsidies can also be justified on redistributive grounds. By allowing low-income households to access quality housing, a more egalitarian distribution of 'necessary commodities' can be achieved (Rosen, 1985).

We can also rationalise the provision of housing subsidies in terms of efficiency considerations. Poor housing quality in low-income neighbourhoods can generate negative externalities such as overcrowding, low residential mobility, or the declining health outcomes of locals (Fisk et al., 2007). By improving the housing quality of low-income recipients, housing subsidies can correct these externalities through an improvement in neighbourhood quality.

3.3.2 The labour supply effects of housing subsidies

Despite the equity and efficiency considerations in favour of the provision of housing subsidies, a common concern for economists and policy makers relates to their impact on the labour supply of claimants. Below, we summarise the theoretical and empirical literature on the likely labour supply impacts of housing subsidies.

3.3.2.1 Theory

Most theoretical assessments of the relationship between housing subsidies and labour supply are based on the neoclassical standard static model of labour supply (for examples, see Moffitt, 2002; Susin, 2005). Here, we outline this model based on the summary provided in Shroder (2010). Consider a scenario where the amount of housing subsidy a claimant is entitled to is determined by the following equation:

$$S = B - T(Y)$$

Where S is the housing subsidy amount, and B is the base subsidy, i.e. the housing subsidy a household with no income is entitled to. $T(Y)$ is what Shroder (2010) defines as the contribution function, that is, the marginal tax rate at which increases in income Y translate into deductions in housing subsidies. The rate at which $T(Y)$ rises, and consequently, housing subsidies fall, in response to a rise in income is called the taper (or withdrawal) rate. In England, for households with below minimum income, $T(Y)$ is zero (see Section 3.2). Above this minimum income there is a taper rate of 65% – housing subsidies are withdrawn at the rate of 65 pence for each pound of additional income. Now let us suppose we have a housing subsidy recipient who is able to work. Her willingness to do so will depend on the opportunity cost of not working (leisure), which depends on earnings from both labour and housing subsidies – the higher the earnings, the higher the opportunity cost, and increased labour income is also punished by reduced housing subsidies. Assume for simplicity that our agent's earnings are going to be spent on one consumption good (housing). Under this scenario, housing subsidies introduce both a substitution and an income effect:

1. The *substitution effect* arises because compared to a no-subsidy scenario, housing subsidies reduce net wages through $T(Y)$, resulting in a lower opportunity cost of leisure. Increased subsidies will lead to higher marginal tax rates because they increase the opportunity cost of labour (by increasing earnings in the absence of labour). In the UK case, due to the high taper rate (65%) the income increases of claimants from increased work effort are 'punished' by a particularly high marginal tax rate. Increased housing subsidies therefore make labour relatively 'costly' in comparison to leisure, inducing a substitution effect towards the latter. Conversely, a reduction in housing subsidies should have the opposite effect: a reduction in the opportunity cost of labour and increased incentives to provide work effort. As it is stated in Shroder (2010), taper rates might also induce extensive margin responses to housing subsidies by changing the slope of claimants' budget constraints (and thereby opportunity costs) to the extent where a corner solution (zero hours of work) becomes the optimal choice.

2. The *income effect* arises because housing subsidies offer a guaranteed level of housing even in the absence of household income. This base level of effective income should also reduce work incentives as a lower earnings target is now sufficient to sustain an 'appropriate' level of household consumption. A housing subsidy cut under this scenario would reduce the size of the income effect – by providing a lower income guarantee, the cut should mean that a larger share of the target income is

now reliant on labour income.

In summary, the neoclassical model predicts that increased housing subsidies will reduce labour supply through both income and substitution effects. Equivalently, we should expect a reduction in housing subsidies to have the opposite effect (increased labour supply).

A limitation of the static labour supply model is that its implications rely heavily on the assumption that housing subsidies supplement income (Shroder, 2010). If instead we treat them as a subsidy on a commodity (that is housing) it is possible that housing subsidies will induce work effort by increasing the demand for the consumption of housing with additional income. Under this scenario, a housing subsidy is effectively a price cut to housing consumption, which will lead to higher real income (Murray, 1980). Higher real income could then lead to increased demand for housing, other goods, and/or leisure. However, whether there is increased demand for leisure, and a corresponding reduction in labour supply, depends on household preferences for the consumption of other goods. Under these circumstances, the labour supply impacts of changing housing subsidies are ambiguous as they depend on heterogeneous household preferences.

An additional theoretical consideration is outlined in Schone (1992). More often than not, housing subsidy programs offer a minimum level of housing consumption. If enrolment in these programs induces claimants to spend more of their other income on housing, they will need to earn more labour income to afford minimum levels of other items (food, utilities, etc.) they consume. Under this scenario, housing subsidies lead to an increase in labour supply, and a reduction in subsidy entitlements should have the opposite effect.

Housing subsidies might also increase the labour supply of claimants by reducing the uncertainty around their housing circumstances and allowing them to spend more time seeking employment (Collinson et al., 2015). Conversely, reduced housing subsidies would then increase housing uncertainty and reduce the amount of time claimants spend seeking work.

Overall, it can be said that economic theory provides ambiguous predictions on the likely effects of housing subsidy changes on the labour supply of claimants. Determining these impacts is therefore mostly an empirical challenge (Jacob and Ludwig, 2012).

3.3.2.2 Empirical evidence

Recent U.S. studies (Mills et al., 2006; and Jacob and Ludwig, 2012) utilise the random selection of housing subsidy recipients to investigate whether housing subsidy provision leads to a reduction in labour supply. Through voucher lotteries that result in the random assignment of housing vouchers to households, these studies estimate the impact of housing assistance in a difference-in-differences (diff-in-diff) framework. Both studies find that housing subsidies had a negative impact on the labour supply of claimants. Non-experimental studies from the U.S. arrive to similar conclusions: Olsen et al. (2005) find that housing assistance in the form of vouchers led to a reduction in household earnings from other sources of at least 30%, while Susin (2005) finds a reduction in household labour earnings of around 10%. These studies however rely on the comparison of claimant and non-claimant groups (the two groups being observed using different data sources in both cases) and are therefore suspect to selection bias due to unobserved differences in behavioural patterns across comparison groups (see Wood et al., 2008).

Also looking at the U.S., a study by Carlson et al. (2012) employs a diff-in-diff design with matched treatment and control households to assess the impacts of housing vouchers on the earnings and labour force participation of Wisconsin residents. They find that housing voucher provision can be associated with a reduction in earnings but has little effect on labour force participation. For Australia, Whelan (2004) finds that housing subsidies lead to reduced labour supply along the extensive margin but finds no evidence of a similar effect along the intensive margin.

For the UK, Gibbons et al. (2018) investigate the impact of the bedroom tax (see Section 3.2) on various household and individual level outcomes in the social rental sector (SRS). They use data from the Understanding Society survey to assess the impacts of the tax on household and individual level outcomes. Their estimates are from a diff-in-diff model in which they compare claimants that were liable to pay 'bedroom tax' to claimants who were not. They find a significant negative effect on the labour income of claimant households and no effect on labour supply along the extensive margin. Interestingly however, the pattern that seems to emerge from their results is that claimants work less in response to subsidy cut. One possible explanation for this is that unlike the LHA cuts, the 'bedroom tax' was not meant to encourage work effort by claimants; the policy was introduced in order to encourage claimants to relocate to smaller sized accommodation. It is evident from the findings in Gibbons et al. (2018) that claimants mostly responded to this policy by moving to smaller (and likely less expensive) flats, offsetting the reduction in disposable income from the subsidy cuts and making a labour supply response obsolete.

Overall, most of the recent studies in the literature on the labour supply impacts of housing subsidies seem to provide evidence in support of the predictions of the neoclassical theory of labour supply: both experimental and non-experimental studies tend to find that housing subsidies have a negative effect on labour supply. The literature however largely focuses on case studies where housing subsidies are introduced, and no study (to our knowledge) focuses explicitly on the labour supply impacts of housing subsidy cuts for existing claimants. We bridge this gap in our empirical analysis below. In addition, there are several methodological challenges in the econometric estimation of the labour supply impacts of subsidy changes that need to be considered.

Shroder (2010) identifies four types of bias that commonly impact on estimates of the labour supply impacts of housing subsidies. First, bias from reporting error, that is, having incorrect information on variables due to incorrectly answered survey questions, can arise from housing subsidy receipts being misreported by survey respondents. Second, selection bias can arise if selection into housing assistance is due to unobserved behavioural patterns among claimants. Comparing claimants to non-claimants therefore carries the risk of (wrongly) attributing differences in labour market outcomes to claimant status, as opposed to differences in other behavioural patterns. Experimental (or quasi-experimental) variation in subsidy provision between (or within) claimant groups provides a good way to control for this issue (see Jacob and Ludwig, 2012). Third, simultaneity bias can arise from the reverse effect of labour supply choices on housing subsidy entitlements. If claimants increase labour supply and subsequently wage income, they may no longer be eligible to receive housing subsidies. According to Shroder (2010) this is relatively easy to control for in a panel setting where housing subsidy entitlements and labour market outcomes can be observed at different points in time: by determining claimant status in earlier time periods we can assess subsequent changes in labour supply even if specific individuals exit the claimant pool. Lastly, omitted variable bias can arise from a variety of other factors impacting labour supply, such as income from other benefits, human capital, or labour market constraints.

3.4 Data

To assess the labour supply impacts of recent PRS housing subsidy cuts in England, we use data on housing benefit claimants from the Understanding Society (US) survey, covering the time period 2009 to 2015. The short sample period is due to the fact that the US survey only started in 2009. The US survey, also known as the UK Household Longitudinal Study (UKHLS), is a longitudinal survey of 40,000 UK households. The US survey combines four different survey samples. First, in the General Population Sample (GPS), households are selected randomly from a clustered sample of UK postcode sectors. Second, the Ethnic Minority Boost Sample (EMBS) is based on an oversampling of the main ethnic minority groups in the UK, so that at least 1,000 households are included in the survey (in wave 1) from these groups. An extension to this sample, the Immigrant and Ethnic Minority Boost sample has also been part of the survey in recent waves. Third, the survey includes a sample of households from the British Household Panel Survey (BHPS), which is the predecessor of the US survey and has been running since 1991. Finally, the Innovation Panel sample conducts some methodological experiments and is therefore not used in the present analysis. Data from the US survey is made available by the UK Data Service. In Petersen et al. (2013), the respondents surveyed in the GPS (which includes most of the households surveyed in US) are found to be representative of the census population at the neighbourhood level.

Our data can be described as an unbalanced panel where individuals are observed in waves (these need not overlap with years). We use data from the first six waves of the US survey. Some individuals are not measured in consecutive waves. We track individuals using the cross-wave person identifier (‘*pidp*’). Adult individuals that share the same household all receive claimant status even if only one of them is indicated as a housing subsidy claimant – whilst from an administrative point of view only one person claims the subsidies, those accrue to the entire household. Under these circumstances, household outcomes are considered, with two individuals forming a benefit unit. When a single person occupies a household, she is the only benefit unit. In the case of large families, all adults form part of a benefit unit and receive claimant status while children do not[71]. Our sample includes housing subsidy claimants from the rental market, and only from England. We exclude SRS tenants with a 'spare room' so that the impacts of the 'bedroom tax' (introduced in 2013) do not bias our estimates (see Section 3.5 below). Overall, our sample contains 10,536 claimants, 31.4% of whom resided in the PRS before the reform and the remaining 68.6% resided in the SRS. Summary statistics on key variables are provided in Appendix B.

3.5 Identification Strategy

In this section, we use individual level panel data from the Understanding Society (US) longitudinal survey to examine the extent to which the LHA cuts impacted the labour market outcomes of housing subsidy claimants living in England.

We estimate the impacts of the housing subsidy cuts using a panel difference-in-differences (diff-in-diff) model. In our diff-in-diff estimation, we compare changes in outcomes for private rental sector (PRS) claimants to changes in outcomes for social rental sector (SRS) claimants, over time. The model is an extension of the specification in Braakmann and McDonald (2018), who estimate the impact of LHA cuts on mobility outcomes. Our baseline specification takes the following form:

$$\mathrm{y_{it}} = \alpha_0 + \theta_i + \theta_t + \theta_w + \theta_{rt} + \sum_{\rho=1}^{7} \gamma_\rho * T_\rho * PRS_i + \sum Z_i^{'} * T_p + \epsilon_{it} \tag{3.1}$$

where:

[71] Naturally, we only look at the labour market outcomes of adult benefit claimants and the children that form a part of their household are not considered in of our analysis.

- $\mathrm{y_{it}}$ are indicators of labour market outcomes (for individual 'i' at time 't') from the US survey, namely: 1) whether an individual is employed full-time; 2) whether an individual is employed part-time; 3) whether an individual participates in the labour market either by working or by actively looking for work; 4) whether an individual is unemployed; 5) the (log) hours of work by those in employment; 6) whether an individual would like to work or not and; 7) whether an individual has a second job or not;

- θ_i, θ_t , θ_w, θ_{rt} are individual, month-year, wave, and region-year fixed effects, respectively;

- PRS_i is an indicator of whether a claimant is in the PRS just before the reforms – this dummy indicates exposure to the reform;

- T_p are dummies for each specific sample year where $\rho = [2009,\ 2010,\ \ldots,\ 2015]$;

- γ_p are the coefficients of interest – they aim to estimate the extent to which outcomes were impacted by the LHA cuts in a given year ρ ;

- γ_p is estimated for each year T_p;

- The term $\sum Z_i' * T_p$ includes individual specific control variables (measured each year) interacted with the treatment period;

3.5.1 Control group specification and identifying assumptions

The sample is restricted to individuals and members of households that claimed housing benefits in the pre-reform (before 2012) period in the social (SRS) or private rental sectors (PRS). The LHA cuts only applied to those renting from the PRS, claimants from the PRS therefore constitute the treatment group: the group 'treated' with the policy change. SRS claimants are untreated by the reform and are therefore included in the control group[72]. The identifying assumption of our diff-in-diff estimation is that, in the absence of the LHA cuts, PRS claimants and SRS claimants would have followed parallel trends in outcome variables (see Angrist and Pischke, 2008). If trends in outcomes change between the two groups after the LHA cuts, we associate these changes with the housing subsidy reductions. We test this assumption using event study plots later in this chapter.

A concern related to our identification strategy is that of endogenous selection into (or out of) the treatment group, due to, for example, individuals moving out of the PRS in response to the reform. To address this, we fix the selection into treatment and control groups in the pre-reform period. This is also important in order to avoid the simultaneity bias mentioned in Section 3.3.2. Since labour market outcomes partly determine claimant status, we must fix selection into the claimant group before we observe changing labour market outcomes as a result of the policy. Otherwise, labour market responses to the reform could remove some individuals from the treatment group.

Summary statistics for our treatment and control groups, before the reform, can be found in Table 3.1. The table summarises mean values for relevant covariates and outcome variables in the pre-announcement period. Balancing tests are also carried out to see whether there are significant pre-treatment differences in some variables. We include other sources of benefit income (for example job seeker's allowance or employment support allowance) as potential covariates to control for the impact of changes in other benefits on our outcomes (see Braakmann and McDonald, 2018).

In general, the pre-treatment differences between the treatment and control groups ob-

[72] Not all SRS claimants were untreated during our sample period due to the introduction of the so-called 'bedroom tax', which applied to SRS claimants with a 'spare room' as of April 2013 (see Gibbons et al., 2018). So that the impact of the bedroom tax does not interfere with our results, individuals/households eligible for this tax are dropped from the sample (it is possible to identify these claimants based on the policy criteria).

Table 3.1: *Summary Statistics – Before LHA Reform Announcement*

Variable	Treatment	Control	Difference-in-means (t-test)
Covariates			
Age	38.84	45.53	6.69***
% student	0.03	0.03	0.00
Children in HH	1.13	0.88	-0.27***
Female	0.66	0.67	0.01
JSA	0.14	0.16	0.01*
ESA	0.02	0.01	0.00
CA	0.03	0.04	0.01*
IBA	0.08	0.14	0.06***
Outcomes			
% would like a job	0.22	0.25	0.03
% full-time	0.12	0.05	-0.07***
% part-time	0.17	0.08	-0.08***
% unemployed	0.24	0.23	-0.01
% participating	0.53	0.34	-0.18***
% having a second job	0.02	0.01	-0.01**
Hours worked	24.74	21.41	-3.33***
Observations	897	2,290	

Notes: JSA stands for jobseeker's allowance, ESA is employment support allowance, CA is carer's allowance and IBA is incapacity benefit allowance.

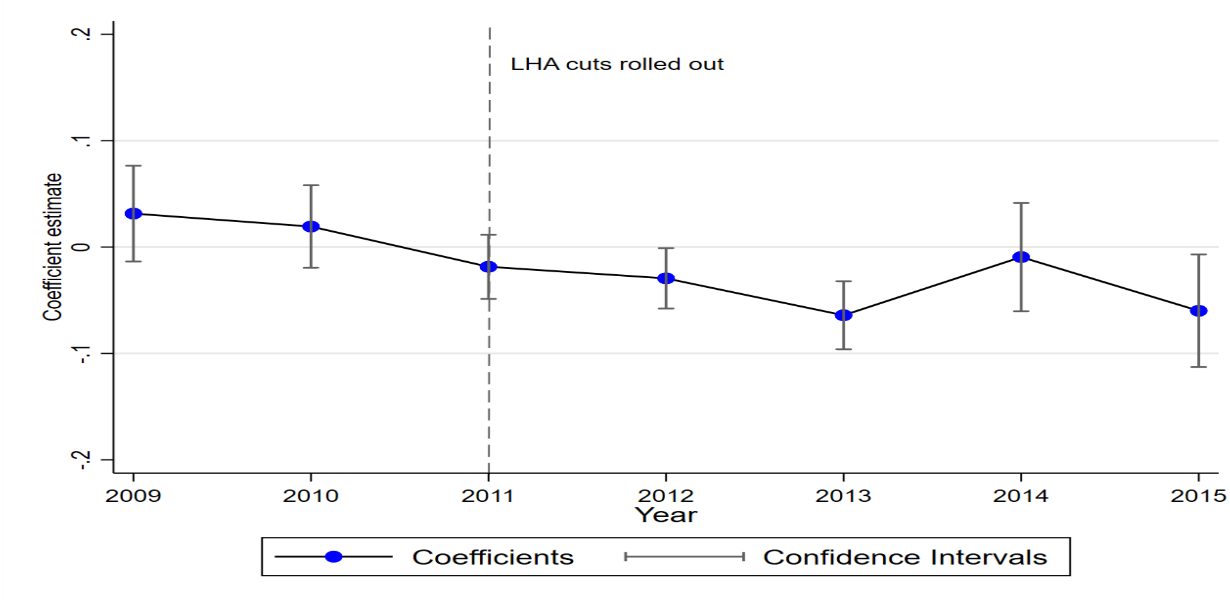
served in Table 3.1 are not a threat to the diff-in-diff identification assumptions (Daw and Hatfield, 2018). On the other hand, confounders that are correlated with both treatment assignment and post-treatment trends in the outcome variable could lead to biased estimates. For example, the SRS control group is on average older in comparison to the treatment group (see Table 3.1). If age specific trends in outcomes coincide with the LHA cuts, we would likely (wrongly) contribute the impact of this to the policy change. To control for this, a common suggestion in the diff-in-diff literature is to interact covariates with the treatment period (see for example Kahn-Lang and Lang, 2019). We do this by picking the covariates that significantly differ between our treatment and control groups (see Table 3.1) and interacting them with the time indicators (T_p) in our estimations[73]. In our empirical model, the covariates used are age, number of children, and a dummy indicating whether the claimant is a recipient of incapacity benefit allowance (IBA). All other possible controls are fixed over time and their effects on the outcome variables should be captured by the individual fixed effects (see Gibbons et al., 2018).

Another important assumption for identification is that, in comparison to the control group, the LHA cuts have resulted in an 'effective' housing subsidy cut for the treatment group. In the presence of rent adjustments, rent affordability might not decline even after a reduction in housing subsidies, prompting no change in individual/household outcomes (see Brewer et al., 2014). To statistically assess whether the LHA reforms can be associated with an effective housing subsidy cut for our treatment group, we estimate Equation 3.1 using the share of rent covered by housing subsidies (subsidy coverage) as the outcome variable. We plot the point estimates corresponding to different sample year 'treatment' effects in Figure 3.1. The negative coefficients shown after the LHA reform roll-out are indicative of an overall negative impact of the cuts on housing subsidy coverage in the PRS (relative to the SRS). Significant point estimates of -0.064 (for 2013), and

[73] We estimated various versions of our empirical model including/excluding interaction terms with control variables. Results are robust to the set of covariates used.

-0.059 (for 2015) indicate that for these two years, the LHA cuts have led to (on average) a roughly 6 percentage point decrease in housing subsidy coverage for claimants in the PRS.

Figure 3.1: *Housing subsidy coverage of rents - event study plot to test impact of LHA cuts on PRS claimants*



Notes: Confidence intervals are drawn for 95% confidence – significance at the 5% level is indicated for each coefficient by the vertical bars (confidence intervals) not spanning zero.

3.5.2 Model specification

To control for the impact of time-fixed individual characteristics on labour market outcomes, we include individual fixed effects. We also include month-year fixed effects to control for cyclical fluctuations in outcomes; wave fixed effects to account for survey wave specific trends and shocks; and region-year fixed effects to control for unobservable time-varying changes at the regional level. We cluster standard errors at the individual level to account for within-individual correlation of error terms over time.

Our specification is a variant of an event study design. In this, we estimate 'treatment' effects ($\gamma_p * T_p * PRS_i$ in equation Equation 3.1) for each year of the sample and then plot the corresponding point estimates. Employing this design is reasonable in our case because of the staggered introduction of the LHA cuts: the reform announcement took place in June 2010, and the transition period commenced in April 2011. In the transition period, claimants were under transitional protection and were only rolled into the new subsidy system in 2012 (see Brewer et al., 2014). Using the event study design, we can track how each of the reform stages affected our outcome variables:

- Assuming no anticipation effects, the first three years are placebo treatments: since no actual change to the LHA system takes place, large and significant coefficients indicate diverging trends between treatment and control groups. The pre-reform period is therefore used to evaluate the parallel trends assumption: that values for the outcome variables would have followed parallel trends in the absence of a treatment (Angrist and Pischke, 2008). If we find diverging trends pre-reform, we cannot validate the parallel trends assumption[74].

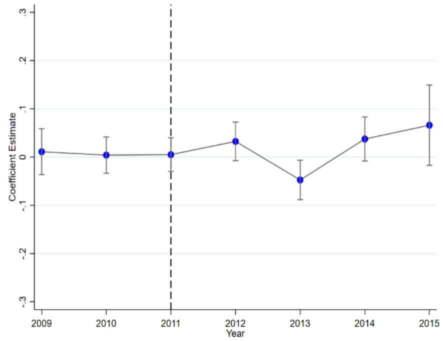
[74] Naturally, this assumption concerns whether trends are parallel post-treatment – as we cannot evaluate this (we do not observe the true counterfactual), we use pre-treatment trends to assess the likelihood that the assumption is going to be valid.

- All subsequent years correspond to point estimates of treatment effects in all subsequent post-reform periods.

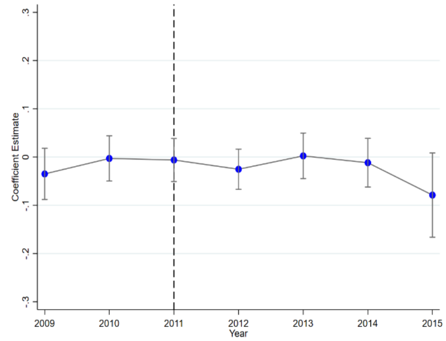
We use the event study plots in Panels (a)-(g) of Figure 3.2 to assess the impact of LHA cuts on claimants' labour market outcomes. In our event study plots, the y-axis shows point estimates for treatment impacts from the LHA cuts in each year of the sample. To find evidence of a significant impact we expect to see: 1) coefficients close to zero and no significant treatment effects in the first three years and 2) coefficients significantly different from zero in subsequent (post-reform) years[75].

[75] In the event study plots, significance at the 5% level is indicated by the confidence intervals (vertical spikes surrounding the coefficients) not spanning zero. We also present the corresponding point estimates in Table B.2 of Appendix B.

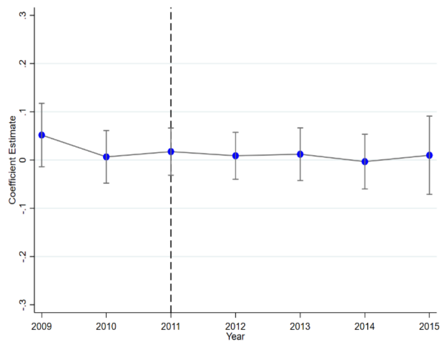
Figure 3.2: *Event Study Plots - Baseline*



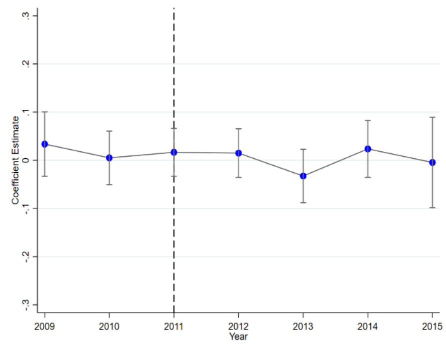
(a) Probability of being employed full-time



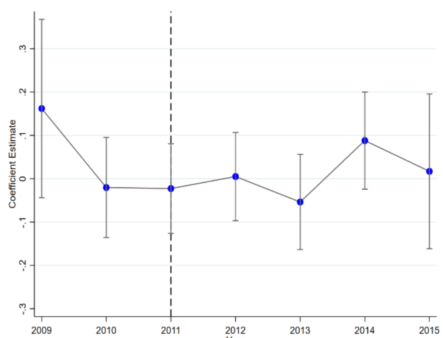
(b) Probability of being employed part-time



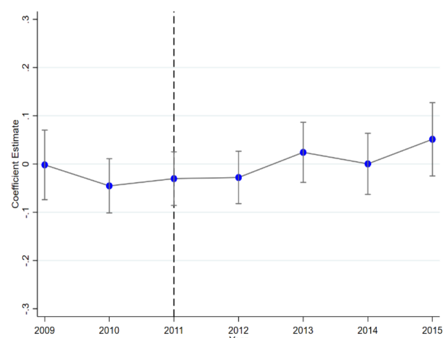
(c) Probability of being unemployed



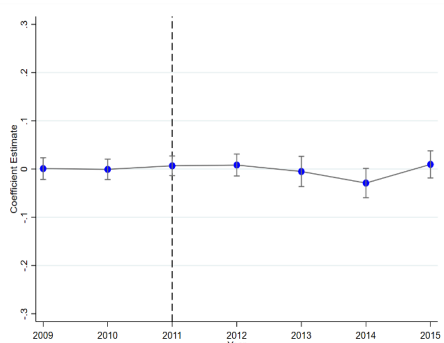
(d) Probability of participation



(e) Hours of work (ln)



(f) Would like to work



(g) Probability of having a second job

3.6 Results

The event study plots (see Figure 3.2) of our baseline estimates do not provide evidence of a significant impact from the LHA cuts on any of the labour market outcomes investigated. Nearly all point estimates are small and insignificant, and there is no discernible post-reform deviation (relative to the control group) in the labour market outcomes of PRS claimants. The only significant point estimate is found for full-time employment, for the year 2013 – the coefficient estimate of -0.0476 indicates that the introduction of LHA cuts in the PRS led to an on average 4.76 percentage point reduction in the probability of having a full-time job, relative to the control group. This result is not supported by point estimates for full-time employment in other post-reform years, as these are positive but imprecisely estimated (see Figure 3.2). Given that all other point estimates are small and random, it is entirely plausible that our finding for full-time employment is a consequence of random noise and is not related to the policy change investigated. To check for the robustness of our findings, and ensure that these are not biased by sample selection or modelling assumptions, we perform the following sensitivity checks.

3.6.1 Robustness checks

To evaluate the plausible robustness of our results, we make several changes to our baseline specification in Equation 3.1. More specifically, we test whether results change under: 1) a specification where we exclude disabled and/or retired individuals; 2) different regional sample specifications; 3) a specification where we restrict our sample to include only women; 4) and a specification where non-claimant PRS residents act as the control group.

3.6.1.1 Excluding disabled/retired individuals

Our baseline sample includes subsidy claimants who are disabled or have already retired from the labour market. It is possible that these individuals are unable to return to the labour market as their personal circumstances might prohibit them from seeking (and engaging in) employment, thereby putting downward bias on our estimates of labour market impacts. We test the robustness of our results to our decision to include disabled/retired individuals by re-estimating our diff-in-diff model using a sample that does not include this group. Roughly 28% of our sample falls within this category, along with nearly 20% of PRS claimants. The results are summarised in the Appendix B, Table B.3. The point estimates presented in Table B.3 provide us with no credible evidence of a change in labour market outcomes in response to the housing subsidy cuts. While we do observe significant point estimates for full-time employment and part-time employment, for both variables the paths of the outcome variable consistently deviate from those of their respective control groups. For example, the positive post-reform deviations for part-time employment are predicated upon positive pre-reform differences (albeit imprecisely estimated) between control and treatment groups, lending no validity to the parallel trends assumption.

3.6.1.2 Regional specifications

Second, our diff-in-diff model from Section 3.5 is estimated for different regional samples in order to assess potential regional heterogeneity in the impacts of the subsidy cuts. Based on previous findings in the literature, we can expect the reform's impact to differ between London and the rest of England (see Brewer et al., 2014). The Shelter (2015) review of the LHA cuts also provides some evidence that London claimants were impacted more severely in comparison to claimants in the rest of England. For this reason, we estimate Equation 3.1 using separate regional samples for London; and for the rest of the England[76]. Results from the regional specifications are summarised in Table B.4 to Table B.5 in the Appendix B. For brevity, we do not show these results on event study plots. Nonetheless, the coefficient estimates in these tables correspond to event study point estimates for each sample year. In line with our baseline estimates, in the regional specifications we find no evidence that the LHA cuts affected the labour market outcomes of claimants.

3.6.1.3 Women only specification

Our baseline estimates were based on a sample that included both male and female claimants. According to the literature on welfare programs, women are more likely to be the target of welfare schemes and are also more sensitive to changes in associated labour supply incentives than men (see Meghir and Phillips, 2010 or Blundell et al., 2016). To take this into account, we re-estimate our diff-in-diff model using a sub-sample that only includes female claimants. Results are summarised in Table B.6 of the Appendix B. The point estimates for the women only sample are mostly small and insignificant providing little evidence that the LHA cuts can be associated with changed labour market outcomes for female claimants. For full-time employment we do observe an initial negative (in 2013) and subsequent positive (in 2014) impact from the subsidy cuts. These effects are difficult to reconcile as they nearly cancel each other out, suggesting that they are more likely to be a consequence of random fluctuations in full-time employment levels than a structural impact of subsidy reform. We also observe a significant positive effect on claimants' willingness to work in 2015, but once again this effect is not supported by the

[76] As a further robustness check, we also estimate a specification where the South East region of England, often thought to follow similar housing market trends as London, is included in the same regional sample with London. This robustness check has no noticeable impact on our results.

path of the outcome variable post-reform. We conjecture that it is unlikely that it took four years for the subsidy cuts to start affecting individual work incentives, and a more likely explanation is that this finding is a consequence of random noise.

3.6.1.4 Alternative control group specification

Finally, to check the robustness of our results to the choice of control group, we re-estimate our model using non-claimants residing in the PRS as the control group. In this specification, outcomes are compared between PRS claimants of housing subsidies and non-claimants, before and after the LHA cuts. Results are summarised in Table B.7 of Appendix B. In Table B.7, we can observe significant post-reform point estimates for both the 'hours of work' and 'would like to work' variables. Note however, that these significant impacts are predicated on significantly different pre-reform trends in both cases, violating the parallel trends assumption and lending no validity to associated findings. In the alternative control group specification, the parallel trends assumption is violated for four of our seven outcome variables, indicating that the non-claimant group is not a reliable comparison group in our case.

3.6.2 Discussion

The results presented in this section provide no credible evidence of an aggregate level impact from the LHA cuts on the labour market outcomes of claimants. Indeed, our null findings suggest that the policy was highly ineffective in its intended objective of encouraging labour supply amongst PRS claimants of housing benefits. For most outcomes, these null effects seem to be fairly precisely estimated, at least in the full sample model, where even the upper limits of confidence intervals would suggest reasonably small (between 4 and 7 percentage point) post-reform changes in labour market outcomes[77].

So what could explain the apparent lack of labour market responses to the subsidy cuts? One explanation is that most subsidy claimants who could (or wanted to) work were already in employment before the reforms to the subsidy system. According to Fenton (2010) the claimants that did not participate in the labour market before the reforms were mostly disabled, sick, or retired. Some claimants also reside in regions where unemployment is pervasive, and long-term structural issues in the labour market may prohibit them from finding employment. Moreover, it is possible that claimants on out-of-work benefits, such as income support or employment and support allowance, are not incentivised to seek employment as doing so would lead to a withdrawal of income from other sources (Shelter, 2015). Finally, while we do find evidence of a significant reduction (on average around six percentage points) in housing subsidy coverage for our sample of claimants (see Section 3.5.1), it is unclear whether this reduction is sizeable enough to trigger a change in aggregate level labour market behaviour. For example, consider a household whose monthly rent is £500. Holding other things constant, for this household our estimates of the reforms' effect on housing subsidy coverage would imply a roughly £30 monthly reduction in disposable income from reduced housing subsidies. Given the high marginal tax rates for claimants who decide to take up work (see Section 3.3.2.1), there might be easier ways for them to make up for this loss in disposable income, for example through reduced consumption or by moving to cheaper rental accommodation. For this reason, it is entirely plausible that similar housing subsidy cuts would have a larger effect on labour market outcomes in tax-benefit systems where marginal tax rates (taper rates) from taking up work are lower. This reflects on a limitation of our findings in that they are expected to have little external validity – the effects of housing subsidy reduction will likely be heterogeneous across different tax-benefit systems that each create a diverse set

[77] An exception is the hours of work outcome, where coefficients tend to be less precisely estimated, most likely due to the low sample size of claimants who are in employment.

of incentives for claimants.

Overall, our results suggest that at the margin, labour market outcomes may not be highly dependent on changes in housing subsidies. From a policy perspective, this indicates that while the government is unable to accomplish their intended objective of bringing claimants back to work, they might force claimants, especially ones already in severe difficulty, under more vulnerable housing conditions. Nonetheless, our present analysis does not quantify these likely mechanisms behind the null effect observed. Consequently, the only inference that we can make based on our results is that, relative to the control group, the reforms have had no significant (and robust) aggregate level impact on the labour market outcomes of PRS claimants.

3.6.3 Limitations

Our analysis has a few limitations that are worth noting here. First, there is considerable attrition present in the US survey, resulting in an unbalanced panel of individuals in our data. For example, note that in our extended sample we have a total of 10,536 observations for 2,734 individuals and 6 waves of the US survey (see Table B.1, Appendix B). This indicates that in our sample, a specific individual is surveyed approximately 3.85 times in 6 waves. It is therefore possible that individuals are only measured once or twice in the period before or after the policy change. This is a limitation because ideally, we would like to see the entire time path for our outcome variables for all individuals so that aggregate level information would reflect annual data on our entire sample[78].

Another limitation is that, in our empirical specifications we do not look at responses

[78] While this leads to a significant reduction in sample size, we did run a robustness check where only those individuals who appear in at least 5 of 6 waves are included. Relevant point estimates under this specification are still not significant, not however that this may also be due to the very low sample size.

to the LHA cuts that might have been overlapping with, or offsetting, labour market responses. For example, if individuals moved to cheaper accommodation in response to the subsidy cuts and managed to increase their disposable income, this likely cancelled out the need for a substantial labour market response. Previous empirical studies on the LHA cuts find mixed evidence of a mobility response (Brewer et al., 2014; Braakmann and McDonald, 2018). In our case, measurement issues related to the US survey render any investigation of mobility outcomes potentially misleading. This is because the variables that can be used to indicate whether individuals moved or not in a given year are featured with very different frequencies in different survey waves, making the construction of time series problematic[79]. Nonetheless, since the explicit aim of the LHA reform was to increase labour supply among claimants, the finding of a null effect for these outcomes is important regardless of what the overlapping mobility response was. Whether the lack of a labour supply response is a consequence of mobility responses or other factors does not change the implication that the policy was ineffective in its main objective of bringing claimants back to work.

Lastly, a limitation of this study is that it is not able to utilise timing differences arising from the roll-out of the reform. Recall from Section 3.2 that the staggered roll-out of the reform resulted in claimants being rolled into the new subsidy system at different times. Unfortunately, our data do not provide information on the date of enrolment into the new subsidy system. Knowing the exact date of enrolment would allow us to define the treatment period for each claimant more precisely instead of relying on aggregate level information at the year level. Future studies on the effects of the subsidy cuts, using more detailed data on enrolment dates, could make use of this timing variation to identify individual or household level impacts more precisely.

[79] These two variables are called 'mvyr' and 'plnowy4'. The latter seems to be used much more often in later survey waves, however using this variable results in a large drop in aggregate observations for mobility in later waves.

3.7 Conclusions

This chapter looked at the labour supply impacts of a reform to the English housing subsidy system that has led to a substantial reduction in subsidy entitlements for private rental sector (PRS) claimants. These subsidy cuts were intended to encourage labour market participation and increased work effort by claimants. To estimate the effects of this policy on the labour market outcomes of claimants, we followed a difference-in-differences approach in which we compared PRS claimants to claimants renting from the social rental sector, who were unaffected by the policy change. Our findings indicate that the reforms had no significant and robust impact on the labour market outcomes of affected claimants, rendering the policy ineffective in its intended objective of bringing claimants back to work. These null findings are robust to a number of different specifications and sensitivity checks. Nonetheless, the precise mechanisms behind our findings are not identified in this chapter, and more detailed analysis is needed to understand the factors that drive the labour market decisions of subsidy claimants.

Bibliography

A. Abadie, A. Diamond, and J. Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, 2010. 20, 22, 29, 32, 39

A. Abadie, A. Diamond, and J. Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015. 22, 30, 32

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, 2008. 118, 122

M. Asplund, R. Friberg, and F. Wilander. Demand and distance: evidence on cross-border shopping. *Journal of Public Economics*, 91(1-2):141–157, 2007. 22

P. Berster, M. Gelhausen, W. Grimme, H. Keimel, S. Maertens, H. Pabst, and D. Wilken. The impacts of the planned air passenger duty in Germany. In *Infraday conference*. TU Berlin, 2010. 23, 24

T. Besley, N. Meads, and P. Surico. The incidence of transaction taxes: Evidence from a stamp duty holiday. *Journal of Public Economics*, 119:61–70, 2014. 53, 60, 73, 97

M. C. Best and H. J. Kleven. Housing market responses to transaction taxes: Evidence from notches and stimulus in the UK. *The Review of Economic Studies*, 85(1):157–193, 2018. 49, 52, 53, 58, 59, 60, 64, 66, 67, 69, 83, 86, 93, 95, 97

A. Billmeier and T. Nannicini. Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3):983–1001, 2013. 22, 31, 32, 33

R. Blundell, M. Costa Dias, C. Meghir, and J. Shaw. Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753, 2016. 127

N. Braakmann and S. McDonald. Housing subsidies and property prices: Evidence from england. *Regional Science and Urban Economics*, 2018. 103, 116, 118, 131

M. Brewer, C. Emmerson, A. Hood, and R. Joyce. *Econometric analysis of the impacts of Local Housing Allowance reforms on existing claimants*. Department for Work and Pensions, 2014. 103, 106, 120, 122, 126, 131

M. Brons, E. Pels, P. Nijkamp, and P. Rietveld. Price elasticities of demand for passenger air travel: a meta-analysis. *Journal of Air Transport Management*, 8(3):165–175, 2002. 25

D. Carlson, R. Haveman, T. Kaplan, and B. Wolfe. Long-term earnings and employment effects of housing voucher receipt. *Journal of Urban Economics*, 71(1):128–150, 2012. 101, 112

M. Chamon, M. Garcia, and L. Souza. Fx interventions in brazil: a synthetic control approach. *Journal of International Economics*, 108:157–168, 2017. 30

R. Chetty, J. N. Friedman, T. Olsen, and L. Pistaferri. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804, 2011. 58, 64

R. Collinson, I. G. Ellen, and J. Ludwig. Low-income housing policy. In *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, pages 59–126. University of Chicago Press, 2015. 101, 108, 111

B. Dachis, G. Duranton, and M. A. Turner. The effects of land transfer taxes on real estate markets: evidence from a natural experiment in toronto. *Journal of Economic Geography*, 12(2):327–354, 2011. 60

J. R. Daw and L. A. Hatfield. Matching and regression to the mean in difference-in-differences analysis. *Health Services Research*, 53(6):4138–4156, 2018. 120

Deutsche Bundesregierung. Bericht an den deutschen bundestag uber die auswirgunken der einfuhrung des luftverkehnrsteuergesetzes auf den luftverkehrssektor und die entwicklung der steuereinnahmen aus der kuftverkehrsteuer., 2012. 20, 23

DW. German air passenger tax under increased fire, 2017. URL http://www.dw.com/en/german-air-passenger-tax-under-increased-fire/a-40168081. 20

EBAA. A snapshot of european aviation taxes, 2015. URL https://www.ebaa.org/european-affairs/taxes-aviation-taxes-europe. 29

Edinburgh Airport. The impact of reducing apd on scotland's airports. Edinburgh Airport, 2015. 20

D. Eiser. A primer on the scottish parliament's new fiscal powers: what are they, how will they work, and what are the challenges? *Fraser of Allander Economic Commentary*, 41(2): 26–41, 2017. 94

European Commission. Housing taxation: from micro design to macro impact. *Quarterly Report on the Euro Area*, 14(1):27–32, 2015. 48

European Commission. Taxation trends in the European Union: Data for the EU Member States, Iceland and Norway. *Publications Office of the European Union*, 2018. 48

European Commission. Taxes in the field of aviation and their impact, 2019. URL https://op.europa.eu/en/publication-detail/-/publication/0b1c6cdd-88d3-11e9-9369-01aa75ed71a1. 24

M. Falk and E. Hagsten. Short-run impact of the flight departure tax on air travel. *International Journal of Tourism Research*, 21(1):37–44, 2019. 21, 26, 31, 34, 42, 43

FCC Aviation. German Aviation Tax Act, 2011. URL `http://www.fccaviation.com/downloads/aviation-tax-act.pdf`. 23

A. Fenton. How will changes to local housing allowance affect low-income tenants in private rented housing? 2010. 107, 129

W. J. Fisk, Q. Lei-Gomez, and M. J. Mendell. Meta-analyses of the associations of respiratory health effects with dampness and mold in homes. *Indoor Air*, 17(4):284–296, 2007. 108

P. Forsyth, L. Dwyer, R. Spurr, and T. Pham. The impacts of Australia's departure tax: Tourism versus the economy? *Tourism Management*, 40:126–136, 2014. 21

C. Fritzsche and L. Vandrei. The German real estate transfer tax: Evidence for single-family home transactions. *Regional Science and Urban Economics*, 74:131–143, 2019. 52, 53, 59

S. Gibbons, M. Sanchez-Vidal, and O. Silva. The bedroom tax. *Regional Science and Urban Economics*, 2018. 106, 113, 118, 120

E. L. Glaeser and E. F. Luttmer. The misallocation of housing under rent control. *American Economic Review*, 93(4):1027–1046, 2003. 97

H. Gordijn and J. Kolkman. Effects of the air passenger tax: Behavioral responses of passengers, airlines and airports. *KiM Netherlands Institute for Transport Policy Analysis: Ministry of Infrastructure and the Environment*, 2011. 21, 26

A. Graham. Demand for leisure air travel and limits to growth. *Journal of Air Transport Management*, 6(2):109–118, 2000. 22

S. Hess and J. W. Polak. Mixed logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management*, 11(2):59–68, 2005. 25, 31, 43

C. A. Hilber and T. Lyytikäinen. Transfer taxes and household mobility: Distortion on the housing or labor market? *Journal of Urban Economics*, 101:57–73, 2017. 48

C. Hofer, R. J. Windle, and M. E. Dresner. Price premiums and low cost carrier competition. *Transportation Research Part E: Logistics and Transportation Review*, 44(5):864–882, 2008. 25

House of Commons, Work and Pensions Committee. Changes to housing benefit announced in the June 2010 Budget, 2010. 101, 107

B. A. Jacob and J. Ludwig. The effects of housing assistance on labor supply: Evidence from a voucher lottery. *American Economic Review*, 102(1):272–304, 2012. 101, 112, 114

J. Jankiewicz and S. Huderek-Glapska. The air transport market in Central and Eastern Europe after a decade of liberalisation–different paths of growth. *Journal of Transport Geography*, 50:45–56, 2016. 22

A. Johannesson-Linden and C. Gayer. Possible reforms of real estate taxation: Criteria for successful policies. *European Economy, Occasional Papers No*, 119(3.3):3–9, 2012. 48

Å. Johansson, C. Heady, J. Arnold, B. Brys, and L. Vartia. Taxation and economic growth. *OECD Working Papers*, 620, 2008. 48

L. Joossens and M. Raw. Smuggling and cross border shopping of tobacco in europe. *BMJ*, 310(6991):1393–1397, 1995. 22

A. Kahn-Lang and K. Lang. The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, pages 1–14, 2019. 120

H. J. Kleven. Bunching. *Annual Review of Economics*, 8:435–464, 2016. 50, 58, 63, 64, 66, 69, 70, 93

H. J. Kleven and M. Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, 128(2):669–723, 2013. 66

A. Knorr and A. Lueg-Arndt. Intercity bus deregulation in Germany–intramodal and intermodal effects after two years. *Research in Transportation Economics*, 59:323–329, 2016. 46

W. Kopczuk and D. Munroe. Mansion tax: the effect of transfer taxes on the residential real estate market. *American economic Journal: economic policy*, 7(2):214–57, 2015. 52, 58, 66, 67, 93

A. Krenek and M. Schratzenstaller. Sustainability-oriented tax-based own resources for the European Union: a European carbon-based flight ticket tax. *Empirica*, 44(4):665–686, 2017. 20, 24, 44

A. Leicester and C. O'Dea. Aviation taxes. *IFS Green Budget*, 2008. 44

K. Mayor and R. S. Tol. The impact of the UK aviation tax on carbon dioxide emissions and visitor numbers. *Transport Policy*, 14(6):507–513, 2007. 21, 44

K. Mayor and R. S. Tol. The impact of European climate change regulations on international tourist markets. *Transportation Research Part D: Transport and Environment*, 15(1): 26–36, 2010. 21

C. Meghir and D. Phillips. Labour supply and taxes. *Dimensions of tax design: The Mirrlees review*, pages 202–74, 2010. 127

G. Mills, D. Gubits, L. Orr, D. Long, J. Feins, B. Kaul, M. Wood, A. Jones, et al. Effects of housing vouchers on welfare families. *Washington, DC: US Department of Housing and*

Urban Development, Office of Policy Development and Research. Retrieved October, 8:2010, 2006. 101, 112

J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba. The Mirrlees Review: conclusions and recommendations for reform. *Fiscal Studies*, 32(3):331–359, 2011. 48

J. A. Mirrlees and S. Adam. *Dimensions of tax design: the Mirrlees Review*. Oxford University Press, 2010. 48, 49, 93, 97

R. A. Moffitt. Welfare programs and labor supply. *Handbook of Public Economics*, 4:2393–2430, 2002. 101, 109

C. Morlotti, M. Cattaneo, P. Malighetti, and R. Redondi. Multi-dimensional price elasticity for leisure and business destinations in the low-cost air transport market: Evidence from easyjet. *Tourism Management*, 61:23–34, 2017. 25

A. Munasib and D. S. Rickman. Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis. *Regional Science and Urban Economics*, 50:1–17, 2015. 30

M. P. Murray. A reinterpretation of the traditional income–leisure model, with application to in-kind subsidy programs. *Journal of public Economics*, 14(1):69–81, 1980. 101, 111

S. B. Nielsen. A simple model of commodity taxation and cross-border shopping. *Scandinavian Journal of Economics*, 103(4):599–623, 2001. 22

E. O. Olsen, C. A. Tyler, J. W. King, and P. E. Carrillo. The effects of different types of housing assistance on earnings and employment. *Cityscape*, pages 163–187, 2005. 112

M. S. A. Ormaechea, M. T. Komatsuzaki, and C. Correa-Caro. *Fiscal reforms, long-term growth and income inequality*. International Monetary Fund, 2017. 33

E. Pels, P. Nijkamp, and P. Rietveld. Equilibrium airfares, frequencies and airport taxes in a multiple airport region: An application of the nested logit demand model. Technical report, Tinbergen Institute Discussion Paper, 1998. 22

J. Petersen, B. Rabe, et al. Understanding society–a geographical profile of respondents. *Understanding Society Working Paper Series No.*, 2013. 115

Revenue Scotland. Annual summary of trends in the devolved taxes. *Revenue Scotland Statistics*, 2017. 96

H. S. Rosen. Housing subsidies: Effects on housing decisions, efficiency, and equity. In *Handbook of Public Economics*, volume 1, pages 375–420. Elsevier, 1985. 108

E. Saez. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212, 2010. 58, 70

A. Sandmo. Optimal taxation: An introduction to the literature. *Journal of Public Economics*, 6(1-2):37–54, 1976. 96

B. S. Schone. Do means tested transfers reduce labor supply? *Economics Letters*, 40(3): 353–357, 1992. 101, 111

Scottish Government. Land and buildings transaction tax (scotland) bill. <https://www.parliament.scot/ResearchBriefingsAndFactsheets/SB_13-02.pdf>, 2013. Last Accessed: 2019-07-15. 49

Scottish Government. Land and buildings transaction tax. <https://news.gov.scot/news/land-and-buildings-transaction-tax>, 2015. Last Accessed: 2019-07-15. 51, 55, 56, 80, 83, 95

N. Seetaram, H. Song, and S. J. Page. Air passenger duty and outbound tourism demand from the United Kingdom. *Journal of Travel Research*, 53(4):476–487, 2014. 21

Shelter. Cuts and changes to housing benefits for private renters (lha), 2015. URL https://england.shelter.org.uk/__data/assets/pdf_file/0020/1142183/7567.01_LHA_Briefing_Report_v4.pdf. 103, 105, 108, 126, 129

M. Shroder. Does housing assistance perversely affect self-sufficiency? a review essay. *Journal of Housing Economics*, 11(4):381–417, 2002. 101

M. D. Shroder. Housing subsidies and work incentives. *Available at SSRN 1691112*, 2010. 102, 109, 110, 111, 114

J. Slemrod, C. Weber, and H. Shan. The behavioral response to housing transfer taxes: Evidence from a notched change in dc policy. *Journal of Urban Economics*, 100:137–153, 2017. 52, 53, 58, 59, 60, 63, 64, 93

U. Steppler. German air travel tax and other duties: A new European trend? *The Air and Space Lawyer*, 24(1):1, 2011. 24

B. Steverink and C. van Daalen. The Dutch taxation on airline tickets. In *the 29th international conference of the System Dynamics Society*, 2011. 22

S. Susin. Longitudinal outcomes of subsidized housing recipients in matched survey and administrative data. *Cityscape*, pages 189–218, 2005. 109, 112

P. Taylor-Gooby. Root and branch restructuring to achieve major cuts: The social policy programme of the 2010 uk coalition government. *Social Policy & Administration*, 46(1): 61–82, 2012. 101

M. H. Thelle and M. la Cour Sonne. Airport competition in Europe. *Journal of Air Transport Management*, 67:232–240, 2018. 25

R. Tunstall et al. The coalition's record on housing: Policy, spending and outcomes 2010-2015. *Social Policy in a Cold Climate Working Paper*, 18, 2015. 106

V. Valdes. Determinants of air travel demand in middle income countries. *Journal of Air Transport Management*, 42:75–84, 2015. 22

S. Whelan. *An analysis of the determinants of the labour market activities of housing assistance recipients*. AHURI, 2004. 113

W. Wilson, C. Barton, and R. Keen. Housing benefit measures announced since 2010, 2016. 102, 105, 106

M. Wood, J. Turnham, and G. Mills. Housing affordability and family well-being: Results from the housing voucher evaluation. *Housing Policy Debate*, 19(2):367–412, 2008. 112

Y. Zhang, M. A. Palma, and Z. P. Xu. Unintended effects of the Alabama HB 56 immigration law on crime: A preliminary analysis. *Economics Letters*, 147:68–71, 2016. 30

J. Zuidberg. The implications of air travel taxes. *Journal of Airport Management*, 10(1): 64–72, 2015. 20, 24, 25, 34, 43

Appendix A

Appendix - Chapter 1

Table A1: *Descriptive Statistics*

Variable	Mean	St. Dev	Min	Max
Passenger numbers (000s)	6656.043	11954.955	24.65	75017.52
Annual % change in passenger numbers	5.47	18.45	-57.78	226.03
Ln (PP per capita)	10.17	0.381	8.748	11.24
Average annual change in ln (PP)	0.207	0.450	-0.192	0.202
Average annual flight ticket price inflation	1.663	1.432	-1.640	12.06

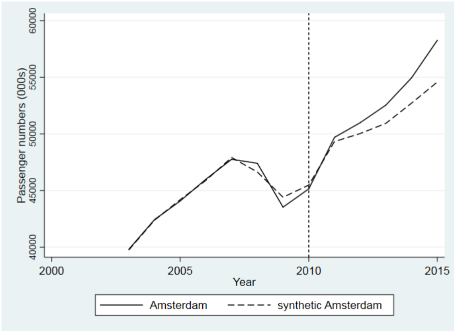
Table A2: *Treated Airports, Control Airports, and Synthetic Control Weights – German Airports*

Treated Airport (Airport Code)	**Control Airports (SC weights)**
Nuremberg (NUE)	Bilbao, Bordeaux, Bratislava, Burgas, Chania, Faro, Gdansk, Girona, Goteborg (0.251), Heraklion, Ibiza, Kerkira, Kos, Larnaka, Malmo (0.242), Marseille, Menorca, Nantes (0.027), Pafos, Porto, Riga, Rodos, Santiago, Sevilla (0.035), Sofia, Stockholm, Stockholm B (0.044), Tallin, Thessaloniki, Toulouse (0.332), Valencia (0.068), Vilnius, Wroclaw
Leipzig (LEJ)	Asturias, Bilbao, Bordeaux, Bratislava, Burgas, Chania, Gdansk, Girona, Jerez, Kerkira, Kos, Larnaka, Lille, Ljubljana, Malmo, Menorca, Montpellier, Murcia, Nantes, Pafos, Reus, Riga, Rodos, Santander (0.037), Santiago, Santorini, Sevilla, Sofia, Stockholm, Stockholm B (0.613), Strasbourg, Tallin, Thessaloniki (0.285), Valencia (0.065), Varna, Vilnius, Wroclaw
Munich (MUC)	Barcelona, Copenhagen, Geneva (0.521), London H, Madrid (0.275), Paris Orly, Palma, Paris CDG (0.152), Stockholm Main (0.052)
Frankfurt Main (FRA)	Athens, Barcelona (0.203), Copenhagen, Helsinki, Lisboa, London H (0.606), Madrid, Paris Orly, Palma, Paris CDG (0.071), Stockholm Main (0.120)
Berlin-Tegel (TXL)	Alicante (0.141), Athens, Barcelona, Copenhagen, Geneva, Helsinki, Lisboa (0.781), Nice, Paris Orly (0.079), Palma, Stockholm Main
Cologne (CGN)	Alicante, Athens, Bordeaux, Budapest (0.417), Faro, Goteborg (0.025), Helsinki (0.216), Heraklion, Ibiza, Larnaka, Lyon, Malaga (0.186), Marseille (0.047), Nice (0.096), Porto (0.013), Riga, Thessaloniki, Toulouse (0.001), Valencia
Dortmund (DTM)	Aalborg, Almeria, Asturias, Bastia, Bilbao (0.113), Bratislava (0.177), Burgas, Chania, Coruna, Gdansk, Granada, Jerez, Kerkira, Kos, Lille, Ljubljana, Malmo, Menorca, Montpellier, Murcia (0.391), Nantes, Pafos, Reus, Rodos, Santiago, Santorini, Sofia, Stockholm, Stockholm B, Tallin, Timisoara, Varna, Vigo, Vilnius (0.318)
Hamburg (HAM)	Alicante, Athens, Budapest (0.116), Copenhagen, Geneva, Helsinki (0.72), Lisboa (0.15), Lyon, Malaga, Marseille, Nice, Paris Orly, Palma (0.014), Porto, Stockholm Main
Dusseldorf (DUS)	Athens, Barcelona, Copenhagen, Geneva (0.082), Helsinki, Lisboa (0.603), Malaga, Nice, Paris Orly (0.314), Palma, Stockholm Main
Munster (FMO)	Aalborg, Almeria (0.131), Asturias, Bastia, Biarritz, Bratislava, Burgas, Chania, Coruna, Granada, Jerez, Kerkira, Kos, Lille, Ljubljana (0.199), Malmo (0.23), Menorca (0.07), Montpellier (0.369), Murcia, Pafos, Reus, Santiago, Santorini, Stockholm, Stockholm B, Tallin, Timisoara, Varna, Vigo, Wroclaw
Bremen (BRE)	Aalborg, Almeria, Asturias, Bastia (0.517), Biarritz, Bilbao, Bordeaux, Bratislava, Burgas, Chania, Coruna, Gdansk, Girona, Goteborg, Granada, Heraklion, Ibiza, Jerez, Kerkira, Kos, Larnaka (0.003), Lille, Ljubljana, Malmo, Menorca, Murcia, Nantes, Pafos, Porto (0.407), Reus, Riga, Rodos, Santorini, Sevilla, Sofia, Stockholm B, Tallin, Thessaloniki, Timisoara, Varna, Vigo, Vilnius, Wroclaw (0.073), Zakintos
Frankfurt Hahn (HHN)	Bilbao, Bordeaux, Bratislava, Burgas, Chania, Faro, Gdansk (0.259), Girona, Goteborg, Heraklion (0.208), Ibiza, Kerkira, Kos, Larnaka, Malmo, Menorca, Nantes, Pafos, Porto, Riga (0.053), Rodos (0.267), Santiago, Sevilla, Sofia, Stockholm, Stockholm B, Tallin, Thessaloniki, Toulouse, Valencia (0.212)
Stuttgart (STR)	Alicante, Athens, Bordeaux, Budapest (0.22), Faro, Goteborg (0.212), Helsinki, Heraklion, Ibiza, Larnaka, Lisboa, Lyon, Malaga (0.448), Marseille, Nice, Porto (0.12), Thessaloniki, Toulouse, Valencia
Paderborn (PAD)	Aalborg, Almeria (0.262), Asturias, Bastia, Biarritz, Bratislava, Burgas, Coruna, Granada, Jerez, Kerkira, Kos, Lille, Ljubljana, Malmo, Montpellier (0.61), Murcia, Pafos, Reus, Santiago, Santorini, Stockholm B, Tallin, Timisoara, Varna, Vigo, Wroclaw, Zakintos (0.128), Zaragoza
Berlin Schonefeld (SXF)	Alicante, Bordeaux, Budapest, Faro, Goteborg, Helsinki, Heraklion, Ibiza, Larnaka, Lyon, Malaga, Marseille, Nantes, Nice, Porto (0.311), Riga (0.568), Rodos, Sevilla (0.121), Thessaloniki, Toulouse, Valencia
Dresden (DRS)	Aalborg, Almeria, Asturias, Bastia (0.051), Biarritz, Bratislava, Burgas, Chania, Coruna, Gdansk (0.065), Granada, Jerez, Kerkira, Kos, Lille, Ljubljana, Malmo (0.109), Menorca (0.152), Montpellier, Murcia, Pafos (0.492), Reus, Santiago (0.09), Santorini, Stockholm (0.032), Stockholm B, Tallin, Timisoara, Varna, Vigo, Vilnius (0.009), Wroclaw, Zakintos
Karlsruhe (FKB)	Aalborg, Almeria, Asturias, Bastia, Biarritz, Bratislava, Burgas, Coruna, Granada, Jerez, Kerkira, Kos, Lille, Ljubljana, Malmo, Montpellier, Murcia, Pafos, Reus, Santiago, Santorini, Stockholm B, Tallin, Timisoara, Varna, Vigo, Wroclaw (0.613), Zakintos, Zaragoza (0.387)
Erfurt (ERF)	Aarhus (0.247), Beziers, Brno, Chios (0.753), Kalamata, Karup, Kavala, Kefallania, Ostrava, Preveza, Samos, Skiathos, Zaragoza
Friedrichshafen (FDH)	Aarhus, Alexandropoulos, Almeria (0.171), Bastia, Biarritz (0.241), Brno (0.022), Coruna, Kavala, Kefallania (0.516), Preveza, Samos, Skiathos, Timisoara (0.05) Zaragoza
Hannover (HAJ)	Alicante (0.3), Bilbao, Bordeaux, Faro, Gdansk, Girona, Goteborg (0.7), Heraklion, Ibiza, Larnaka, Lyon, Marseille, Nantes, Porto, Riga, Sevilla, Sofia, Thessaloniki, Toulouse, Valencia, Vilnius
Saarbrucken (SCN)	Aarhus (0.2692), Beziers (0.308), Brno, Chios, Kalamata, Karup, Kavala, Kefallania (0.4228), Ostrava, Preveza, Samos, Skiathos, Zaragoza

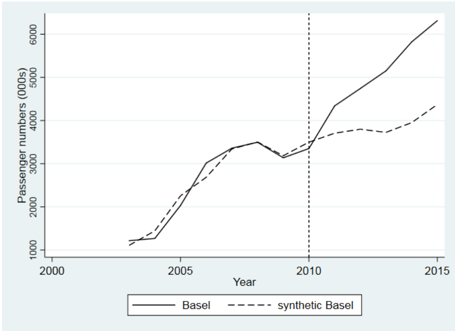
Table A3: *Treated Airports, Control Airports, and Synthetic Control Weights – Bordering Airports*

Treated Airport	**Control Airports (SC weights)**
Amsterdam (AMS)	Barcelona (0.441), Copenhagen, Geneva, Helsinki, Lisboa, London H (0.319), Madrid, Palma, Paris Orly, Paris CDG (0.178), Stockholm Main (0.062)
Eindhoven (EIN)	Bilbao, Bordeaux, Bratislava, Burgas, Chania, Faro, Gdansk, Goteborg, Heraklion, Ibiza, Kerkira, Kos, Larnaka, Lyon, Marseille, Menorca, Nantes, Pafos, Porto, Riga, Rodos, Santiago (0.067), Sevilla, Sofia, Stockholm (0.253), Stockholm B, Tallin, Thessaloniki, Toulouse, Valencia, Vilnius, Wroclaw (0.68)
Rotterdam (RTM)	Aalborg (0.209), Almeria, Asturias, Bastia, Biarritz, Bratislava, Burgas, Chania, Coruna, Granada, Jerez, Kerkira, Kos, Lille, Ljubljana, Malmo, Menorca, Montpellier, Murcia, Pafos, Reus, Santiago (0.053), Santorini, Stockholm, Stockholm B, Tallin, Timisoara, Varna, Vigo, Wroclaw, Zakintos (0.737)
Maastricht (MST)	Aarhus, Alexandropoulos, Beziers (0.254), Brno, Chios, Kalamata, Karup, Kavala, Kefallania, Nimes (0.746), Ostrava, Preveza, Samos, Skiathos, Zaragoza
Basel (BSL)	Alicante, Bilbao, Bordeaux, Budapest, Faro, Gdansk (0.51), Goteborg, Heraklion, Ibiza, Larnaka, Lyon, Marseille, Nantes, Nice, Porto, Prague, Riga (0.153), Rodos, Sevilla, Sofia, Thessaloniki, Toulouse, Valencia (0.337), Vilnius
Luxembourg (LUX)	Aalborg, Asturias, Bilbao, Bordeaux (0.024), Bratislava, Burgas, Chania, Gdansk, Granada, Jerez, Kerkira, Kos (0.071), Lille, Ljubljana, Malmo, Menorca, Montpellier (0.692), Murcia, Nantes, Pafos, Reus, Riga, Rodos, Santiago, Santorini, Sevilla, Sofia, Stockholm (0.213), Stockholm B, Tallin, Varna, Vigo, Vilnius, Wroclaw
Billund (BLL)	Asturias, Bilbao, Bordeaux, Bratislava, Burgas, Chania, Gdansk, Granada, Jerez, Kerkira, Kos, Larnaka (0.056), Lille, Ljubljana, Malmo, Menorca, Montpellier (0.532), Murcia, Nantes, Pafos, Reus, Riga, Rodos, Santiago (0.163), Santorini, Sevilla, Sofia, Stockholm (0.11), Stockholm B, Tallin, Thessaloniki, Varna, Vilnius (0.139), Wroclaw
Brussels (BRU)	Athens, Barcelona, Copenhagen, Geneva (0.209), Helsinki, Lisboa, Malaga, Nice (0.437), Paris Orly (0.355), Palma, Stockholm M
Zurich (ZRH)	Athens, Barcelona, Copenhagen (0.39), Geneva, Helsinki, Lisboa (0.436), Madrid (0.174), Malaga, Paris Orly, Palma, Stockholm Main
Metz (ETZ)	Aarhus (0.284), Alexandropoulos, Beziers, Brno, Chios, Esbjerg (0.628), Kalamata, Karpathos, Karup, Kavala, Kefallinia, Ostrava, Preveza, Samos (0.088), Skiathos
Szczecin (SZZ)	Aarhus, Alexandropoulos, Beziers (0.469), Brno (0.531), Chambery, Chios, Kalamata, Karpathos, Karup, Kavala, Kefallinia, La Rochelle, Nimes, Preveza, Samos, Skiathos
Prague (PRG)	Alicante, Athens, Budapest (0.637), Faro, Geneva (0.363), Helsinki, Ibiza, Lisboa, Lyon, Malaga, Marseille, Nice, Palma, Porto, Stockholm Main
Charleroi (CRL)	Aalborg, Almeria, Asturias, Bastia (0.34), Biarritz, Bordeaux, Bratislava, Burgas, Chania, Coruna, Faro (0.005), Gdansk (0.218), Girona, Goteborg, Granada, Heraklion (0.176), Jerez, Kerkira, Kos, Larnaka, Lille, Ljubljana, Malmo, Menorca, Murcia, Pafos, Porto (0.529), Reus, Riga (0.13), Rodos, Santiago, Sevilla, Sofia, Stockholm B, Tallin, Thessaloniki, Varna, Vigo, Vilnius, Wroclaw

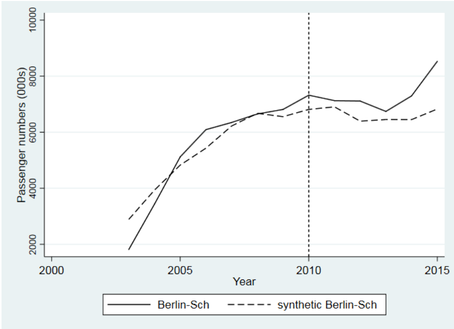
Figure A1: *Synthetic control estimates - Amsterdam to Charleroi*



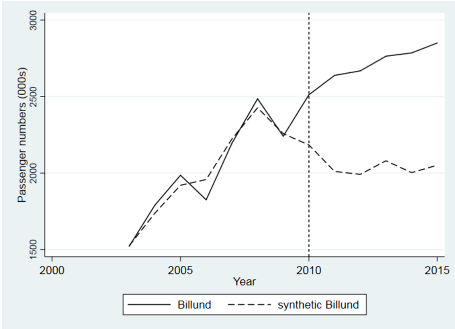
(a) Amsterdam



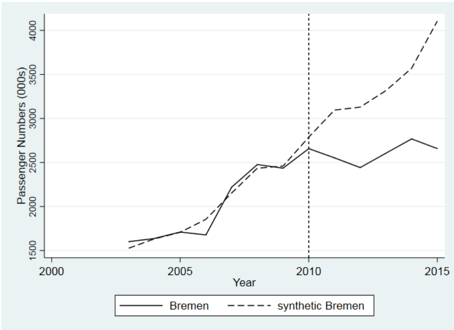
(b) Basel



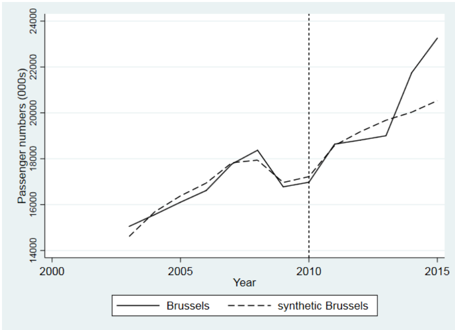
(c) Berlin Schonefeld



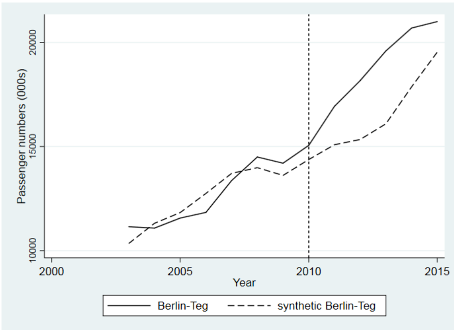
(d) Billund



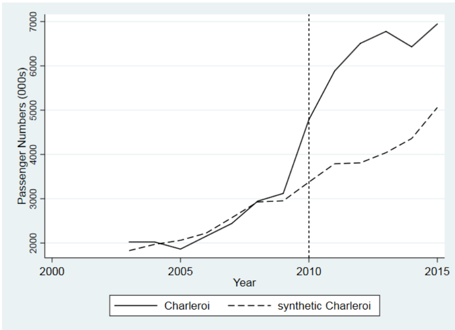
(e) Bremen



(f) Brussels

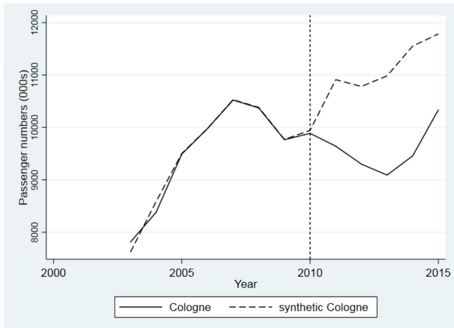


(g) Berlin - Tegel

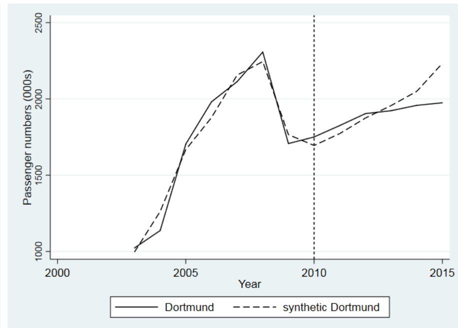


(h) Charleroi

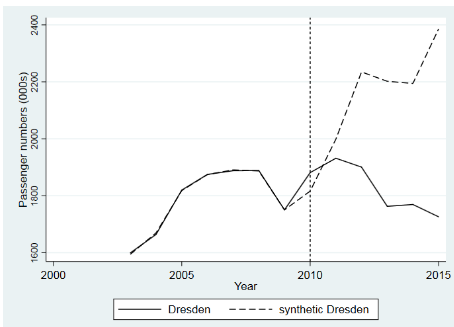
Figure A2: *Synthetic control estimates - Cologne to Frankfurt Main*



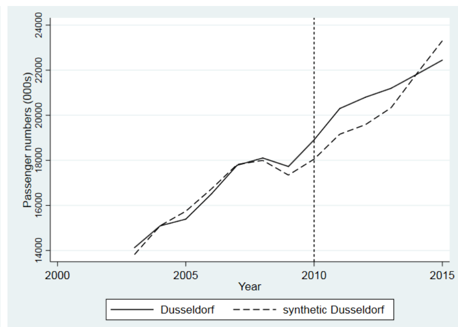
(a) Cologne



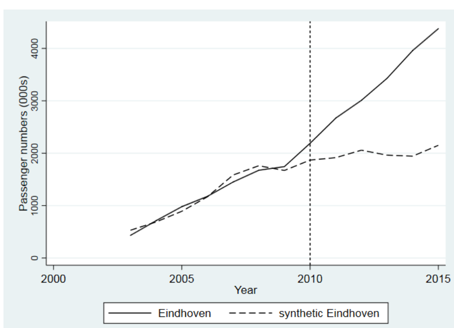
(b) Dortmund



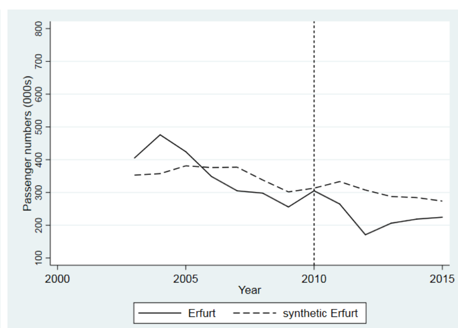
(c) Dresden



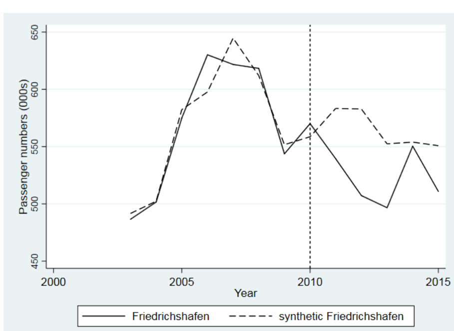
(d) Dusseldorf



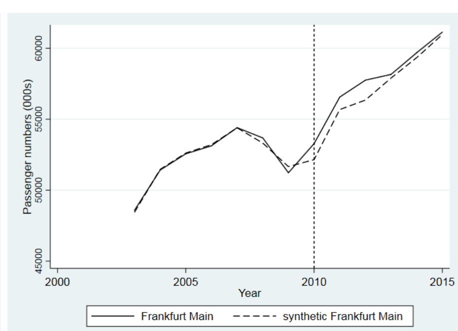
(e) Eindhoven



(f) Erfurt

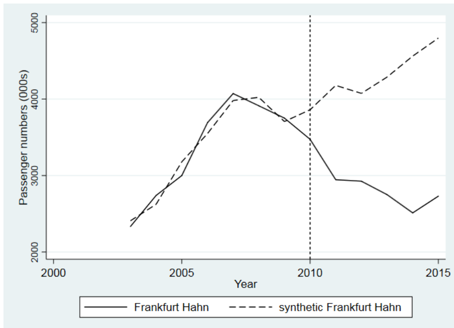


(g) Friedrichshafen

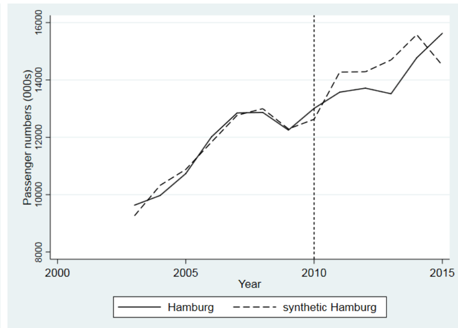


(h) Frankfurt Main

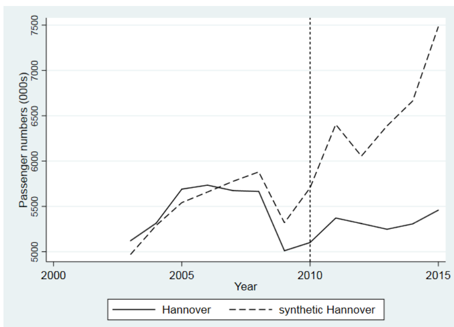
Figure A3: *Synthetic control estimates - Frankfurt Hahn to Metz*



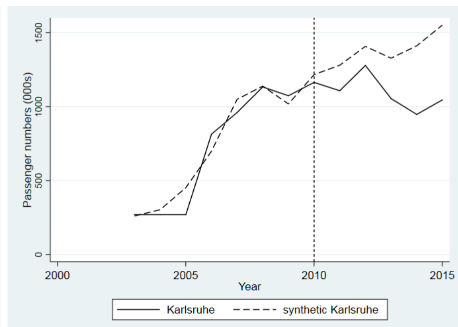
(a) Frankfurt Hahn



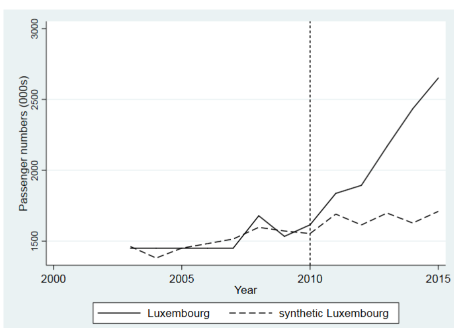
(b) Hamburg



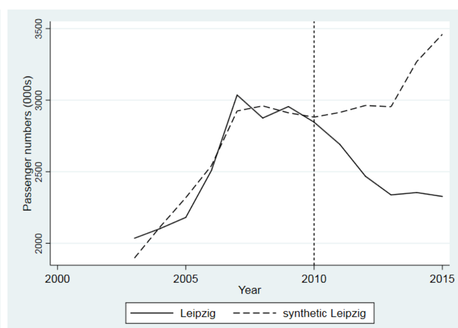
(c) Hannover



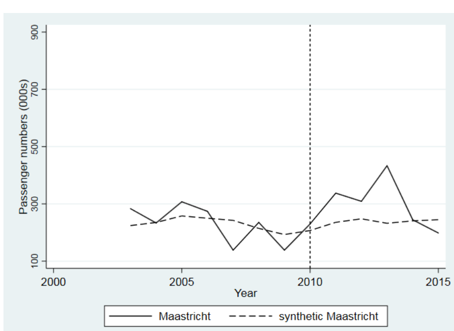
(d) Karlsruhe



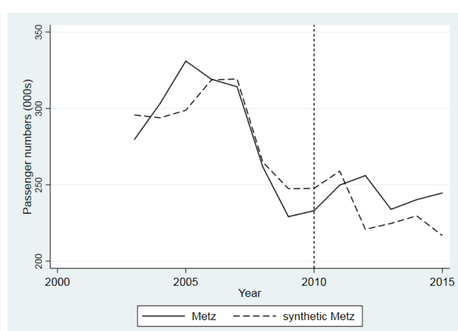
(e) Luxembourg



(f) Leipzig

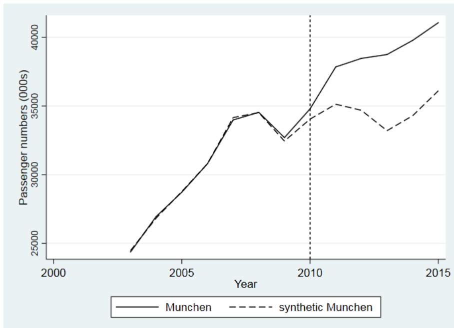


(g) Maastricht

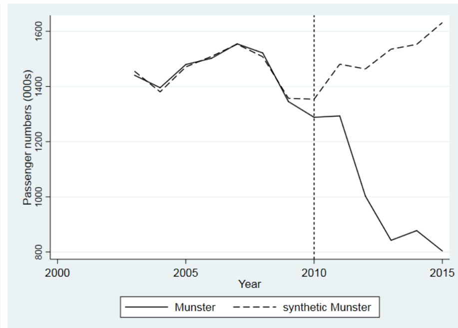


(h) Metz

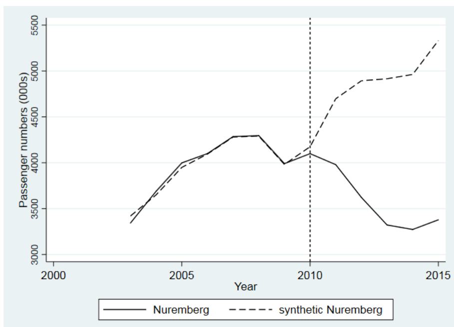
Figure A4: *Synthetic control estimates - Munich to Stuttgart*



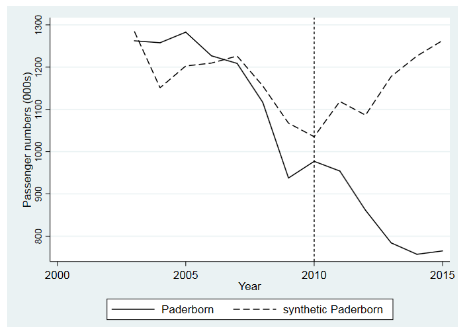
(a) Munich



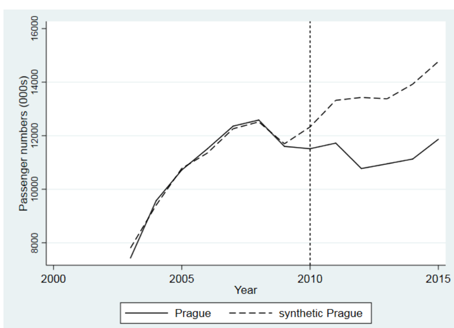
(b) Munster



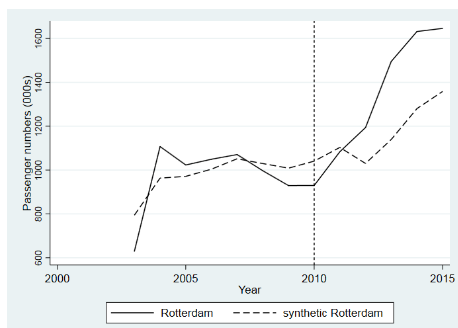
(c) Nuremberg



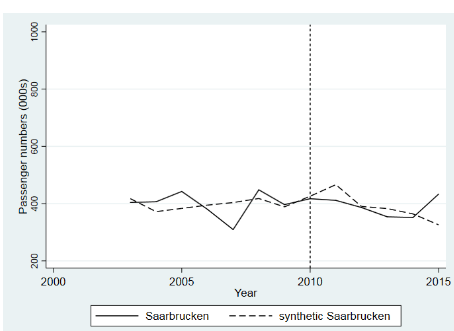
(d) Paderborn



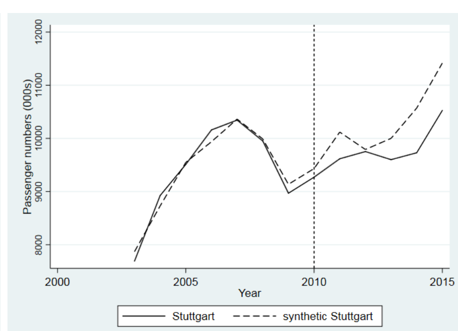
(e) Prague



(f) Rotterdam

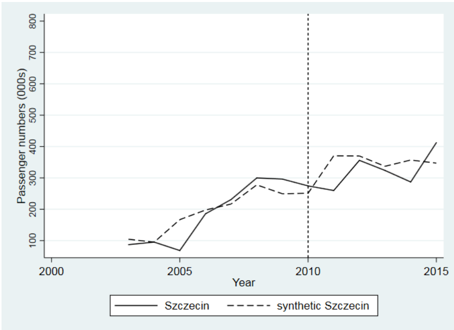


(g) Saarbrucken

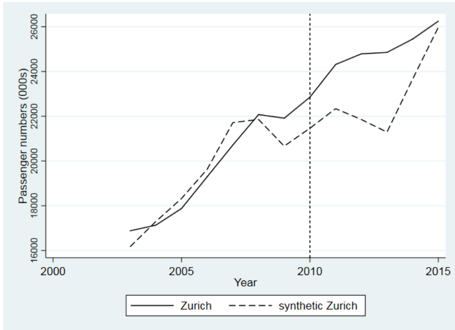


(h) Stuttgart

Figure A5: *Synthetic control estimates - Szczecin and Zurich*



(a) Szczecin



(b) Zurich

Appendix B

Appendix - Chapter 3

Table B.1: *Summary Statistics – Baseline Sample*

Variable	Mean	SD	Count
Assignment			
% PRS (pre-reform)	0.31	0.46	10,536
Covariates			
Age	46.12	15.46	10,536
% student	0.02	0.14	10,536
% retired	0.20	0.40	10,536
% disabled	0.18	0.39	10,536
Children in HH	0.92	1.20	10,536
Female	0.68	0.46	10,536
JSA	0.11	0.32	10,536
ESA	0.02	0.14	10,536
CA	0.04	0.20	10,536
IBA	0.12	0.33	10,536
Outcomes			
% would like a job	0.20	0.40	10,536
% full-time	0.09	0.29	10,536
% part-time	0.13	0.33	10,536
% unemployed	0.19	0.39	10,536
% participating	0.41	0.49	10,536
% having a second job	0.02	0.13	10,536
Hours worked	25.02	11.63	2,013

Table B.2: OLS Results – Baseline sample

Variables	**FT**	**PT**	**UN**	**PR**	**WL**	**SJ**	**HW**
$\gamma_p * PRS_i * 2009$	0.0110	-0.0350	0.0518	0.0337	-0.0018	0.0009	0.1620
	(0.0243)	(0.0271)	(0.0335)	(0.0341)	(0.0368)	(0.0115)	(0.1050)
$\gamma_p * PRS_i * 2010$	0.0040	-0.0029	0.0066	0.0051	-0.0453	-0.0006	-0.0204
	(0.0192)	(0.0239)	(0.0278)	(0.0283)	(0.0287)	(0.0108)	(0.0589)
$\gamma_p * PRS_i * 2011$	0.0050	-0.0061	0.0175	0.0166	-0.0303	0.0068	-0.0227
	(0.0178)	(0.0228)	(0.0250)	(0.0253)	(0.0284)	(0.0104)	(0.0528)
$\gamma_p * PRS_i * 2012$	0.0324	-0.0253	0.0089	0.0150	-0.0279	0.0083	0.0051
	(0.0203)	(0.0213)	(0.0248)	(0.0257)	(0.0277)	(0.0116)	(0.0519)
$\gamma_p * PRS_i * 2013$	-0.0476**	0.0024	0.0121	-0.0326	0.0242	-0.0051	-0.0538
	(0.0209)	(0.0241)	(0.0279)	(0.0283)	(0.0318)	(0.0160)	(0.0560)
$\gamma_p * PRS_i * 2014$	0.0374	-0.0118	-0.0033	0.0237	0.0003	-0.0290*	0.0880
	(0.0233)	(0.0258)	(0.0290)	(0.0301)	(0.0323)	(0.0155)	(0.0572)
$\gamma_p * PRS_i * 2015$	0.0660	-0.0789*	0.0099	-0.0045	0.0513	0.0096	0.0170
	(0.0425)	(0.0446)	(0.0414)	(0.0478)	(0.0388)	(0.0144)	(0.0911)
Observations	10,536	10,536	10,536	10,536	10,536	10,536	2,013
R-squared	0.072	0.039	0.044	0.041	0.081	0.035	0.225
Nr. of individuals	2,734	2,734	2,734	2,734	2,734	2,734	716
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.3: OLS Results – Excluded disabled/retired sample

Variables	**FT**	**PT**	**UN**	**PR**	**WL**	**SJ**	**HW**
$\gamma_p * PRS_i * 2009$	-0.0557	0.0794	0.0186	0.0511	-0.0604	0.0125	0.139
	(0.0538)	(0.0600)	(0.0587)	(0.0601)	(0.0488)	(0.0198)	(0.1350)
$\gamma_p * PRS_i * 2010$	-0.0718	0.110*	-0.0130	0.0219	-0.108***	0.0099	-0.0568
	(0.0507)	(0.0571)	(0.0535)	(0.0554)	(0.0393)	(0.0195)	(0.0999)
$\gamma_p * PRS_i * 2011$	-0.0714	0.111*	-0.0027	0.0374	-0.0805**	0.0133	-0.0696
	(0.0507)	(0.0575)	(0.0533)	(0.0545)	(0.0404)	(0.0195)	(0.0978)
$\gamma_p * PRS_i * 2012$	-0.0749	0.115**	-0.0234	0.0167	-0.0406	0.0047	-0.0317
	(0.0510)	(0.0580)	(0.0532)	(0.0542)	(0.0405)	(0.0189)	(0.0927)
$\gamma_p * PRS_i * 2013$	-0.0436	0.0837	-0.0317	0.0084	-0.0732*	0.0185	-0.0452
	(0.0512)	(0.0563)	(0.0508)	(0.0532)	(0.0411)	(0.0201)	(0.0930)
$\gamma_p * PRS_i * 2014$	-0.109**	0.0928	-0.0044	-0.0192	-0.0404	0.0158	-0.105
	(0.0506)	(0.0576)	(0.0533)	(0.0565)	(0.0412)	(0.0219)	(0.0892)
$\gamma_p * PRS_i * 2015$	-0.0715	0.0695	0.0070	0.0069	0.0628	0.0191	0.0181
	(0.0509)	(0.0542)	(0.0487)	(0.0546)	(0.0417)	(0.0180)	(0.0895)
Observations	7,653	7,653	7,653	7,653	7,653	7,653	2,000
R-squared	0.091	0.052	0.061	0.070	0.081	0.049	0.231
Nr. of individuals	2,210	2,210	2,210	2,210	2,210	2,210	716
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: OLS Results – London specification

Variables	**FT**	**PT**	**UN**	**PR**	**WL**	**SJ**	**HW**
$\gamma_p * PRS_i * 2009$	-0.0316	0.0111	-0.0147	-0.0369	0.0935	0.0074	-0.109
	(0.0552)	(0.0614)	(0.0842)	(0.0802)	(0.0870)	(0.0149)	(0.150)
$\gamma_p * PRS_i * 2010$	-0.0416	-0.0001	-0.0288	-0.0717	0.0285	0.0198	-0.255*
	(0.0495)	(0.0553)	(0.0629)	(0.0592)	(0.0708)	(0.0182)	(0.131)
$\gamma_p * PRS_i * 2011$	-0.0147	-0.0306	-0.0075	-0.0535	0.102	0.0201	-0.136
	(0.0463)	(0.0549)	(0.0581)	(0.0559)	(0.0722)	(0.0189)	(0.148)
$\gamma_p * PRS_i * 2012$	-0.0197	-0.0065	-0.0160	-0.0426	0.0995	-0.0224	0.0307
	(0.0503)	(0.0465)	(0.0617)	(0.0618)	(0.0669)	(0.0137)	(0.145)
$\gamma_p * PRS_i * 2013$	-0.0611	-0.0203	0.0881	0.0122	0.0205	0.0274	-0.127
	(0.0475)	(0.0613)	(0.0724)	(0.0779)	(0.0852)	(0.0280)	(0.141)
$\gamma_p * PRS_i * 2014$	0.0361	-0.0323	-0.0439	-0.0436	-0.0958	0.0051	0.0542
	(0.0568)	(0.0695)	(0.0727)	(0.0826)	(0.0846)	(0.0287)	(0.191)
$\gamma_p * PRS_i * 2015$	0.0796	-0.0656	-0.0140	-0.0009	0.0382	0.0284	0.173
	(0.0808)	(0.0842)	(0.0818)	(0.113)	(0.0849)	(0.0249)	(0.251)
Observations	2,457	2,457	2,457	2,457	2,457	2,457	488
R-squared	0.107	0.084	0.087	0.076	0.115	0.081	0.446
Nr. of individuals	690	690	690	690	690	690	178
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: OLS Results – Rest of England specification

Variables	FT	PT	UN	PR	WL	SJ	HW
$\gamma_p * PRS_i * 2009$	0.0194	-0.0432	0.0712**	0.0551	-0.0223	-0.0017	0.152
	(0.0266)	(0.0305)	(0.0363)	(0.0373)	(0.0408)	(0.0141)	(0.129)
$\gamma_p * PRS_i * 2010$	0.0143	-0.0034	0.0182	0.0264	-0.0627**	-0.0075	0.0004
	(0.0207)	(0.0265)	(0.0309)	(0.0318)	(0.0315)	(0.0128)	(0.0680)
$\gamma_p * PRS_i * 2011$	0.0040	0.0057	0.0255	0.0356	-0.0640**	0.0027	-0.0065
	(0.0189)	(0.0248)	(0.0276)	(0.0279)	(0.0308)	(0.0121)	(0.0530)
$\gamma_p * PRS_i * 2012$	0.0434*	-0.0272	0.0188	0.0339	-0.0544*	0.0138	0.0356
	(0.0221)	(0.0242)	(0.0266)	(0.0278)	(0.0304)	(0.0142)	(0.0580)
$\gamma_p * PRS_i * 2013$	-0.0391*	0.0094	-0.0148	-0.0449	0.0244	-0.0124	-0.0282
	(0.0236)	(0.0256)	(0.0295)	(0.0289)	(0.0334)	(0.0191)	(0.0590)
$\gamma_p * PRS_i * 2014$	0.0266	-0.0024	0.0175	0.0425	0.0197	-0.0358*	0.0272
	(0.0257)	(0.0272)	(0.0312)	(0.0321)	(0.0346)	(0.0184)	(0.0548)
$\gamma_p * PRS_i * 2015$	0.0700	-0.0919	0.0227	0.0005	0.0671*	-0.0045	-0.0970
	(0.0512)	(0.0565)	(0.0486)	(0.0548)	(0.0402)	(0.0173)	(0.0990)
Observations	8,079	8,079	8,079	8,079	8,079	8,079	1,525
R-squared	0.078	0.047	0.054	0.050	0.089	0.037	0.298
Nr. of individuals	2,055	2,055	2,055	2,055	2,055	2,055	539
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: OLS Results – Women only specification

Variables	FT	PT	UN	PR	WL	SJ	HW
$\gamma_p * PRS_i * 2009$	0.0428	-0.0577	0.0262	0.0188	-0.0348	-0.0101	0.222
	(0.0284)	(0.0372)	(0.0415)	(0.0467)	(0.0478)	(0.0122)	(0.150)
$\gamma_p * PRS_i * 2010$	0.0123	-0.0133	0.0105	0.0085	-0.0617*	0.0049	0.0151
	(0.0221)	(0.0325)	(0.0343)	(0.0376)	(0.0366)	(0.0124)	(0.0668)
$\gamma_p * PRS_i * 2011$	0.0005	-0.0034	0.0208	0.0182	-0.0400	0.0015	-0.0465
	(0.0212)	(0.0301)	(0.0318)	(0.0350)	(0.0358)	(0.0115)	(0.0662)
$\gamma_p * PRS_i * 2012$	0.0343	-0.0375	0.0297	0.0261	-0.0365	0.0072	-0.0295
	(0.0238)	(0.0289)	(0.0305)	(0.0335)	(0.0358)	(0.0151)	(0.0634)
$\gamma_p * PRS_i * 2013$	-0.071***	0.0260	-0.0181	-0.0652*	0.0019	0.0079	-0.0360
	(0.0250)	(0.0326)	(0.0328)	(0.0363)	(0.0398)	(0.0205)	(0.0679)
$\gamma_p * PRS_i * 2014$	0.0539**	-0.0226	0.0231	0.0592	0.0006	-0.0343	0.0998
	(0.0272)	(0.0346)	(0.0358)	(0.0389)	(0.0403)	(0.0212)	(0.0655)
$\gamma_p * PRS_i * 2015$	0.0514	-0.0782	0.0330	0.0039	0.0927**	-0.0020	-0.0585
	(0.0576)	(0.0633)	(0.0496)	(0.0624)	(0.0452)	(0.0183)	(0.113)
Observations	7,164	7,164	7,164	7,164	7,164	7,164	1,468
R-squared	0.079	0.058	0.045	0.057	0.095	0.044	0.265
Nr. of individuals	1,847	1,847	1,847	1,847	1,847	1,847	519
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.7: OLS Results – Non-claimant control group

Variables	**FT**	**PT**	**UN**	**PR**	**WL**	**SJ**	**HW**
$\gamma_p * PRS_i * 2009$	-0.183***	0.0156	0.0648	-0.103*	0.152***	0.0104	-0.229**
	(0.0632)	(0.0710)	(0.0518)	(0.0607)	(0.0424)	(0.0243)	(0.0971)
$\gamma_p * PRS_i * 2010$	-0.214***	0.0582	0.0915*	-0.0622	0.0609*	0.0002	-0.271***
	(0.0601)	(0.0696)	(0.0490)	(0.0577)	(0.0347)	(0.0238)	(0.0836)
$\gamma_p * PRS_i * 2011$	-0.193***	0.0618	0.115**	-0.0164	0.0949***	-0.0082	-0.248***
	(0.0606)	(0.0694)	(0.0487)	(0.0564)	(0.0350)	(0.0237)	(0.0844)
$\gamma_p * PRS_i * 2012$	-0.158***	0.0786	0.0573	-0.0190	0.0799**	-0.0044	-0.185**
	(0.0601)	(0.0692)	(0.0488)	(0.0567)	(0.0344)	(0.0231)	(0.0823)
$\gamma_p * PRS_i * 2013$	-0.120*	0.0558	0.0523	-0.00935	0.0397	0.0313	-0.169**
	(0.0612)	(0.0672)	(0.0491)	(0.0584)	(0.0355)	(0.0239)	(0.0849)
$\gamma_p * PRS_i * 2014$	-0.130**	0.0971	0.0234	-0.0129	0.0632*	0.0285	-0.132
	(0.0615)	(0.0695)	(0.0487)	(0.0585)	(0.0351)	(0.0259)	(0.0810)
$\gamma_p * PRS_i * 2015$	-0.111*	0.0652	0.0271	-0.0184	-0.0313	0.0038	-0.138*
	(0.0612)	(0.0659)	(0.0458)	(0.0575)	(0.0357)	(0.0210)	(0.0812)
Observations	12,039	12,039	12,039	12,039	12,039	12,039	7,319
R-squared	0.059	0.030	0.080	0.057	0.046	0.033	0.090
Nr. of individuals	3,954	3,954	3,954	3,954	3,954	3,954	2,652
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at the individual level and are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.