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

Francesco Nicolai
The London School of Economics

Marco Pelosi
The London School of Economics

Simona Risteska *
The London School of Economics

March 28, 2019 – Version 0
Very Preliminary

Abstract

Taxes that happen concurrently with the purchase are more salient than deferred taxes. Using a sharp geographical discontinuity between London Boroughs, we show that the incidence of property taxes deferred to the future is too small compared to the incidence of stamp duty taxes happening at the moment of buying the property. The difference in incidence implies very large discount rates that cannot be easily rationalized even after accounting for liquidity constraints. The lack of salience at the moment of purchase implies that the burden of the tax will be borne in the future to meet the budget constraint. This implies that there is an optimal tax mix, even though one of the two taxes is more distortionary at the moment of purchase than the other.

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

1 Introduction

A standard tenet of economic theory is that the statutory incidence of taxes is irrelevant for their economic incidence[1]. It should also be the case that whether a tax is paid at the moment of transaction or later is irrelevant for its incidence, as long as we take into account the time value of money and the riskiness of the cash flows. By looking at the UK residential property market, this paper shows that this is not the case and that deferred taxes have a markedly lower incidence compared to taxes paid at the time of decision-making.

Together with France, the United Kingdom is one of the few countries receiving a sizeable fraction of revenues from property taxes, amounting to about 4.3% of GDP or more than $£84$ billion in 2016 (European Commission (2018)). The two main taxes levied on domestic properties are the Stamp Duty Land Tax and the council tax. The former is a tax levied on the transaction value of land and any buildings and structures thereon. The fact that its statutory incidence falls on the buyer, who is required to pay the tax liability to the HM Revenue and Customs within very few weeks from the completion of the transaction, and the fact that the tax represents a lump sum ranging between 1% and 7% of the property value are features that make the stamp duty tax particularly salient at the moment of purchase. The latter, which will be the focus of the present paper, is a tax levied by the local government on a yearly basis. The council tax is levied on the resident, as opposed to the house owner, and is based on the property value in 1991. While the council tax is extremely salient at the moment when it needs to be paid, we will show that this is not the case at the moment when properties are purchased even though, in present value terms, it is similar to or even larger than the stamp duty tax. By using the geographical discontinuity at the border of different local authorities in the London area, we are able to estimate the incidence of the council tax on property prices and contrast it with the incidence of the stamp duty tax estimated, among others, by Best and Kleven (2018). The London area is particularly suitable for the estimation because of the sharp nature of the council borders and the large dispersion in council tax rates across councils. For instance, Figure 1 depicts a road that is at the border of the Borough of Kensington and Chelsea with Westminster. As can be seen from the picture, the houses on both sides

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

Figure 1 : A Typical Border


.

of the street are otherwise identical except for the fact that they pay quite different council tax amounts: the ones on the left pay $£2,279$ per year in council tax while those on the right pay $£1,421$ per year. If we discount the future payments as a perpetuity at a rate of 4%, similar to the mortgage rates observed in sample, we obtain that the difference between the two present values amounts to $£21,450$ (about $\$28,000$). The borders are made even sharper by the fact that many London Boroughs share services, such as waste management, and that many other services, such as access to parks, schooling and religious facilities, are not strictly limited to residents of a given Borough. We will show in Sections 4 and 5.1 that the estimated incidence of council tax on property prices is too low even after accounting for time value of money and the fact that discount rates might be larger because of borrowing constraints. This suggests that the stamp duty tax is more distortionary compared to the council tax. On the other hand, the fact that the council tax becomes very salient ex-post implies that it distorts consumption choices as it significantly reduces the available income. In Section 5.2 we will show that this could potentially allow the Government to optimally tune the taxes to minimise distortions for a given level of revenue.

The present paper adds on the burgeoning literature on behavioural public finance and the salience of taxes (or the lack thereof). Chetty et al. (2009) is the first paper to empirically estimate how differences in salience can alter the behaviour of economic agents. They intervene in a grocery store in order to modify the salience of sales taxes and show that the incidence on buyers is largely reduced when taxes are made fully salient. In a second experiment they compare the effect of excises taxes, which are included in posted prices, and sales taxes, which are not explicitly included, on alcohol demand and again show that tax salience plays an important role in consumer behaviour. The setting in the present paper is quite similar to the second experiment in Chetty et al. (2009), given that the stamp duty tax is paid upfront while the council tax is deferred and thus less salient. For policy reasons, however, the question of property taxes is of greater importance because of the large amounts of money involved and the fact that it is very difficult for agents to learn since buying a new property is typically a once in a lifetime event. Following Chetty et al. (2009), other papers have also explored the question of tax salience, for instance, Feldman and Ruffle (2015) and Feldman et al. (2018) have replicated the findings in laboratory experiments, while Finkelstein (2009) similarly shows that the introduction of electronic toll payments raises toll expenditures. Taubinsky and Rees-Jones (2018) further explore the topic by showing that there is large variation in the way agents react to tax salience and investigate policy implications. The present paper is also similar to Allcott (2011) who demonstrates that a similar bias is present in the automobile market, namely, car buyers fail to correctly price in the future energy cost at the time of purchase. As in Allcott (2011), our conclusions also rely on the choice of an appropriate discount factor. We will show in Section 5.1 that the bias persists even after allowing for large discount rates. Finally, the paper extends the literature on property taxes; among others, we will use the results of Besley et al. (2014) and, in particular, Best and Kleven (2018) to compare our estimates of the council tax incidence with the stamp duty incidence estimates in order to highlight the lack of salience of the former.

The rest of the paper is organised as follows: Section 2 describes the data and the institutional setting; Section 3 gives evidence of the geographical distribution of council taxes and points out that this can significantly bias our estimates if not appropriately taken care off, before proceeding with the details of our identification strategies; Section 4 presents the empirical estimates of the council tax incidence; Section 5 develops a stylised model to help interpret the findings and to show that the estimated incidence is too low to be consistent with fully-salient taxes, and explores some policy implications; and finally, Section 6 summarises and concludes the paper.

2 Data

To estimate the incidence of council taxes we need access to data on property characteristics and house prices, as well as council taxes paid. Price paid data on house transactions are readily available from the HM Land Registry website. This dataset contains information about all residential properties transacted in England and Wales between 1995 and 2018 that have been sold for full market value[2]. The dataset comprises of the price paid, the transaction date and, most importantly, the address of the house which allows us to pinpoint the exact location of every property. Additionally, the data provide us with information on the property type, which can be one of five possible categories (a detached, semi-detached, or terraced house, a flat/maisonette and other), the age of the property (classified into old or new to distinguish between newly built properties and already established buildings) and the duration of tenure, i.e., whether the property is under a freehold or leasehold[3].

Since we would ideally like to compare properties that are as similar to each other as possible, we need more information on property characteristics. For this purpose we make use of two additional datasets: the Zoopla Property data and Domestic Energy Performance Certificates. The Zoopla Property data[4] has been collected by Zoopla, one of the UK's leading providers of property data for consumers and property professionals, giving free access to information on 27,000,000 homes, up to 1,000,000 property listings and 15 years of sold prices data. The dataset covers the period between 1st January 2010 and 31st March 2019 for properties located in Great Britain (England, Wales, Scotland).

[2]Data excluded from the dataset include commercial transactions, property transactions that have not been lodged in with HM Land Registry and transactions made below market value. For more details on the property sales not included in the dataset the reader can visit the HM Land Registry website.

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

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

The dataset contains details on characteristics such as property location, property type[5], whether the property has been categorised as residential or commercial[6], number of bedrooms, number of floors, number of bathrooms, number of receptions and whether the property is listed for sale or for rent[7]. In addition, we also have access to the asking price for both rents and sales, however, we use the more accurate transaction price from the HM Land Registry dataset. The second source of house characteristics comes from the Ministry of Housing, Communities and Local Government. On their website, one can access the Energy Performance Certificates (EPC) for domestic and non-domestic buildings. For domestic properties, before 2008 certificates could be lodged on a voluntary basis. From 2008 onwards, however, it has become mandatory for accredited energy assessors to lodge the energy certificates. Consequently, the data coverage drastically improves around that time, as does our ability to match these data with the price paid data. More specifically, the matching rate jumps from about 50 percent to over 90 percent around 2008. The dataset contains information on the location, property type, total floor area, number of storeys, number of rooms, floor level and height, along with many indicators of energy efficiency and quality of glazed surfaces. The final piece of data needed to conduct our analysis is related to council tax data; in the following section we are going to describe in more detail the functioning of this property tax and the relevant data.

2.1 Council Tax

The taxation of properties in the United Kingdom is peculiar compared to other OECD countries, representing a rather large source of both central Government and local authorities' revenues. The three main taxes levied on properties are the council tax, business rates and stamp duty taxes. Council taxes are levied on each occupier, rather than on the owner, of domestic properties. The tax is one of the few levies in Great Britain being both set and collected by local authorities (Boroughs in the case of London) and it repre-

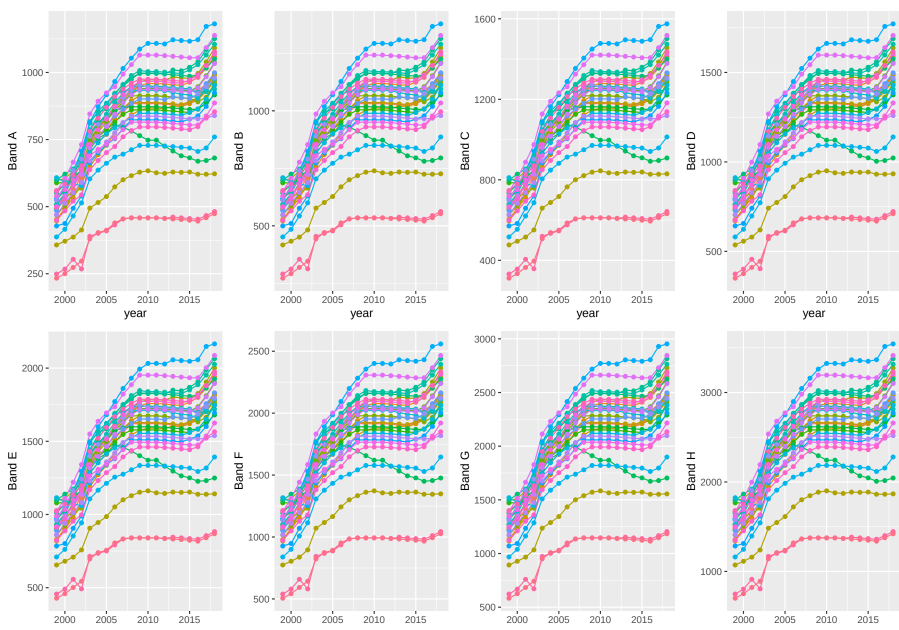
[5]Property types include: barn conversion, block of flats, bungalow, business park, chalet, château, cottage, country house, detached bungalow, detached house, end terrace house, equestrian property, farm, farm house, finca, flat, hotel/guest house, houseboat, industrial, land, leisure/hospitality,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

[6]We keep only properties categorised as residential.

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

sents one of their major sources of revenue (around one-third of total revenue), the other sources being commercial property taxes (business rates) and transfers from the central Government. The tax is based on a classification in eight bands (A-H) based on the value of the property as established by the Valuation Office in 1991; newly built properties are assigned to a band, after having converted their current value into the value of an equivalent property in 1991. Each London Borough is responsible to set each year the annual tax amount to be paid by a property in band D; the amount to be paid by other bands is automatically set as a ratio to the amount in band D[8]. Bands C and D represent the largest fraction of dwellings (about 50 percent of the total), but there is variation across Boroughs with central properties being skewed towards higher valued bands compared to properties in outer Boroughs. Figure 2 shows the time series of the average council tax

Figure 2 : Time Series of Council Taxes



payable per Borough. Each panel in the Figure depicts the average amount payable by different bands showing that, by construction, the tax moves in locksteps across bands. More interestingly, it should be noted that there is a wide dispersion in amounts payable across Boroughs, even though the ranking across different local authorities remains al-

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

most constant with the only exception of the Borough of Hammersmith and Fulham where taxes have been slashed starting from the late 2000s. After a marked increase in council tax rates in the early 2000s, the freeze mandated by the central Goverment after the 2008 financial crisis is visible in the time series; since 2011, taxes can be raised only by a centrally set amount unless a local referendum allows the authority to do so. We will show in Section 3.1 that the geographical distribution of council tax rates is not random and could severely bias any estimate of incidence, given that central (and pricier) Boroughs tend to set lower council tax rates. This is mainly because central Boroughs tend to have larger fraction of properties in higher bands; for instance, the Borough of Kensington and Chelsea raises more than fifty percent of its revenues from bands G and H, while Barking and Dagenham raise less than five percent from such bands.

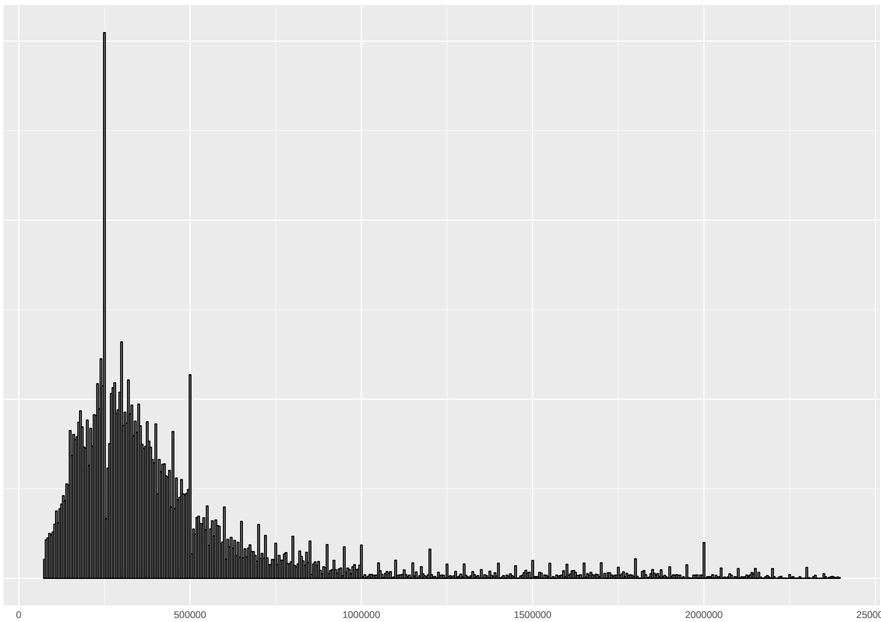
We obtain information on council tax band assignment from the website of the Valuation Office Agency, which provides data on the full address and the council tax band for each property in Great Britain. The average amount to be paid in each London Borough by each band in the period 1999-2018 is obtained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

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

2.2 Descriptive Statistics

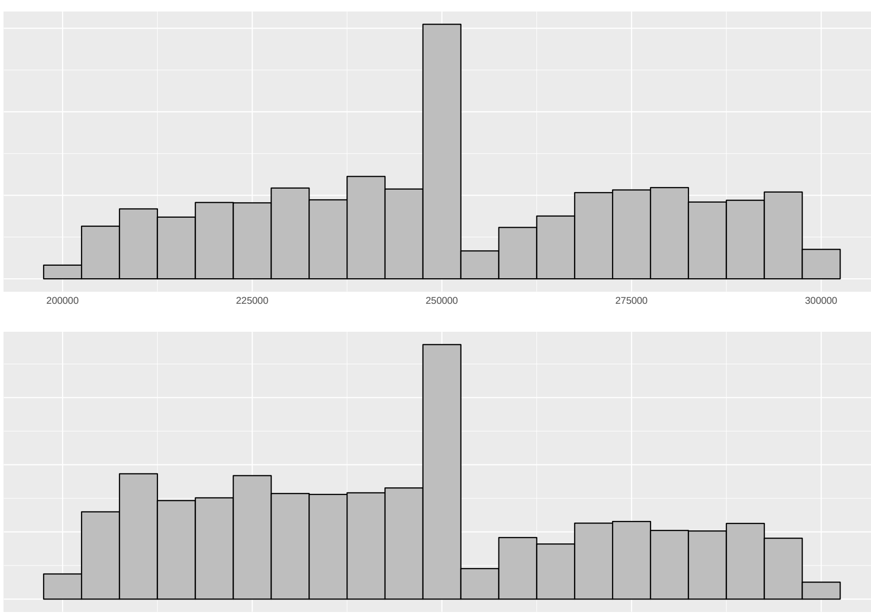
Figure 3 shows the distibution of transaction prices for domestic properties in London, truncated to exclude extremely high property prices which are, however, included in the analysis. It is immediately obvious that there is a large degree of bunching in prices, as noted for instance in Best and Kleven (2018). The bunching mainly happens just before stamp duty notches, which allows Best and Kleven (2018) to estimate the local incidence of this tax. Figure 4, for instance, shows the large extent of bunching at the threshold of $£250,000$ (upper panel) and $£500,000$ (lower panel) where the stamp duty tax jumps from 1% to 3% and from 3% to 4%, respectively. Best and Kleven (2018) estimate a rather large incidence of stamp duty tax on property prices and argue in favour of evidence of rather strict borrowing constraints; we will use their estimates to inform our estimates of the incidence of the council tax, letting us disentangle how much of the incidence is due to borrowing constraints and how much is due to pure time discount. Figure 5 shows

Figure 3 : Histogram of Property Prices in London



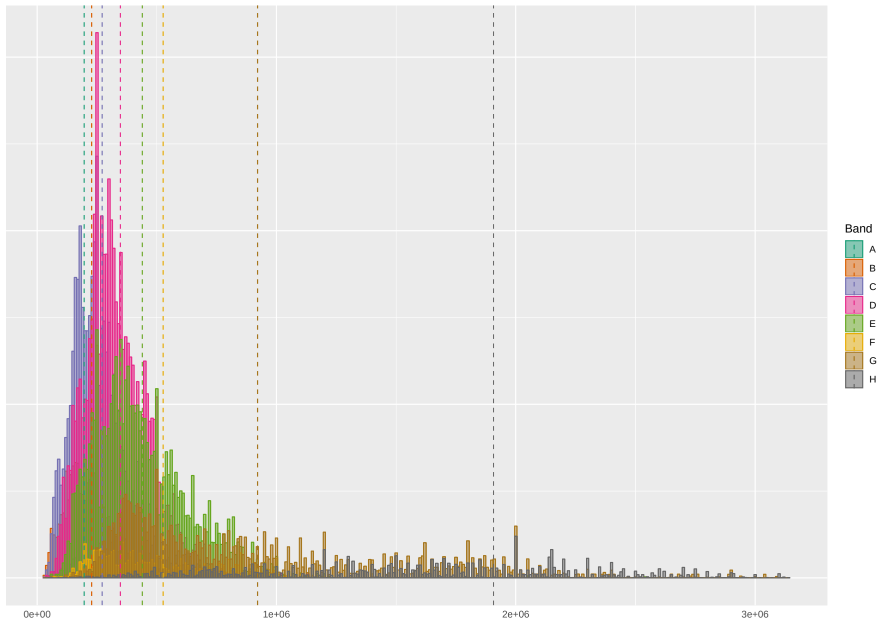
.

Figure 4 : Bunching at Stamp Duty notches



.

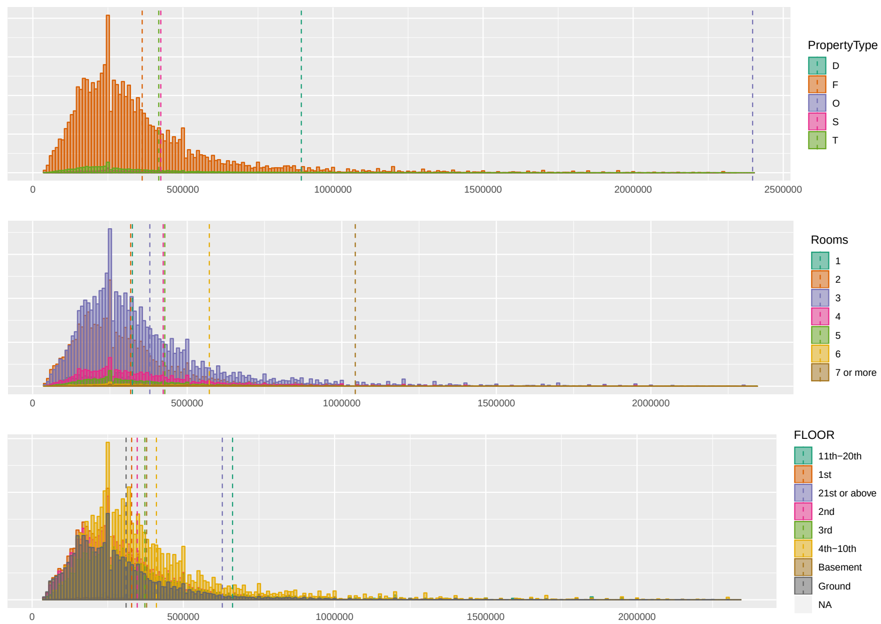
Figure 5 : Histogram of Prices by Band



.

the distribution of house prices per band. As one should expect, higher bands tend to have houses with higher average prices although there is a large dispersion within bands. This is because prices have increased a lot over the past twenty years, especially for more central and higher-banded properties. This makes it essential that we compare only transactions occurring in close periods. Moreover, one should notice that the number of properties belonging to bands C and D dominates the rest, as previously mentioned. Figure 6 shows that there is a wide dispersion of transaction prices depending on house characteristics. The upper panel shows the distribution of prices by property type; the middle panel by number of rooms; the lower panel by floor level. The vertical lines in each panel show the average values for each category. As the first panel shows, there is a disproportionate amount of flats in our sample, which we see as an advantage in our estimation, as flats are much more likely to be similar to each other relative to other property types. If we exclude the residual category (Others), detached houses are most expensive, with an average price of almost £1 million, followed by semi-detached houses and terraced houses, and finally, flats are the cheapest category. In the second panel, one can see that, naturally, the house price is increasing in the number of rooms, however, this relationship is not linear. For example, properties with one and two rooms

Figure 6 : Histogram of Prices by Characteristics



.

have similar average prices, while the price of having three rooms jumps quite significantly. There is also a very large jump in prices between houses with five and six rooms. Finally, the third panel shows that there is no systematic relationship between the floor on which the property is located and its price. The issue with this classification, however, is that flats are over-represented in our sample and this is what drives the price of top floors up, as these are the flats most sought after in taller buildings.

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

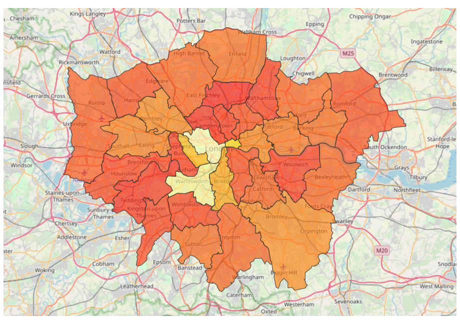
3 Empirical Strategy

3.1 Evidence of Selection

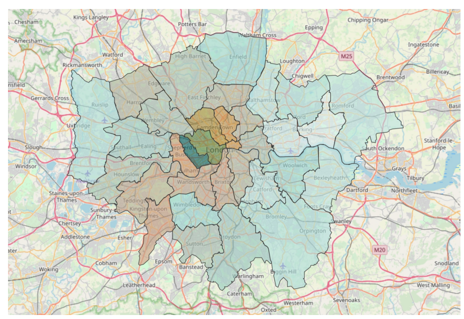
The main issue arising when estimating the incidence of council taxes is the fact that the cross-sectional distribution across councils is strongly correlated with other characteristics that affect house prices. To see this, Figure 7 shows a map of the distribution of Band D average taxes for each London Borough along with the respective distribution

Figure 7 : Council Taxes and House Prices

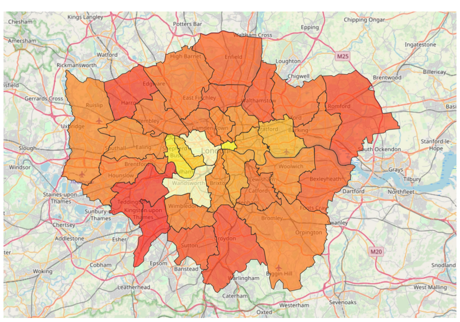
The maps show the distribution of Band D average taxes for each London Borough along with the respective distribution of house prices in 2000 and 2018.



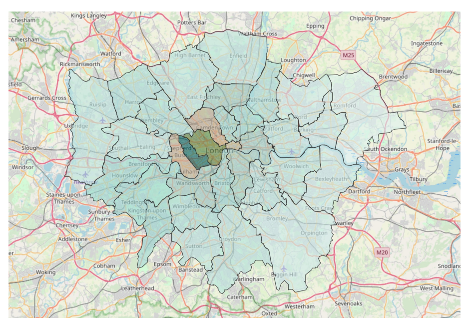
(a) : Council Taxes in 2000



(b) : House Prices in 2000



(c) : Council Taxes in 2018



(d) : House Prices in 2018

of house prices. Panel 7a shows the distribution of council taxes in 2000, where taxes increase moving from yellow to red; Panel 7b the distribution of house prices in the same year, where prices increase moving from light blue to brown. Panel 7c shows the distribution of council taxes in 2018 while panel 7d the distribution of house prices in the same year. It is visually striking that councils with higher taxes tend to have lower prices. For instance, the City of Westminster had the lowest Band D council tax in 2000 ($£375.17$) and the second highest median house price ($£296,250$), after the Borough of Kensington and Chelsea ($£374,975$) which had the fourth lowest council tax ($£623.38$). In 2018 the same holds true, with the City of Westminster having the lowest council tax ($£710.50$) and the second highest median price ($£1,547,000$), after Kensington and Chelsea ($£1,995,000$) which had the fifth lowest council tax ($£1,139.41$). In general, it is clear from the map that Boroughs that lie further from the centre tend to have higher council taxes and lower prices, while the more central Boroughs tend to exhibit the opposite characteristics. To confirm the intuition obtained from Figure 7, we can run a naive regression of median house prices on average house characteristics per Borough, i.e.,:

$$\overline{p}_{dbt} = \alpha + \beta\tau_{dbt} + \gamma\overline{Size}_{dbt} + \delta_{bt} + \zeta\overline{x}_{dbt} + \varepsilon_{dbt} \tag{1}$$

where $\overline{p}_{dbt}$ is the median price of a house in council d, band b at time t; τ_{dbt} is the average council tax for a house in council d, band b at time t; $\overline{Size}_{dbt}$ is the median size for a house in council d, band b at time t; δ_{bt} are band-time fixed effects; and $\overline{x}_{dbt}$ are council-band-time controls which include the average number of rooms, the average lighting, heating and water costs.

Table 1 reports the results of regression (1); the first column provides the baseline results including band-time fixed effects only to remove the mechanical correlation between increasing property prices and taxes over time and the fact that moving from band A to band H goes hand in hand with houses of improving quality. It is remarkable that this naive strategy would imply an enormous and statistically significant incidence, with a point estimate of -316.61 and a t-statistics of 18.42. To give an intuition, using a discount factor of $r = 4\%$ (similar to the one observed in sample) this would roughly imply that an extra $£1$ in the present value of taxes would imply a difference in prices of $r \times \beta =$

Table 1 : Evidence of Selection

The table shows the estimates of a simple regression of council median house prices on median council taxes per borough, namely: $\overline{p}_{dbt} = \alpha + \beta\tau_{dbt} + \gamma\overline{Size}_{dbt} + \delta_{bt} + \zeta\overline{x}_{dbt} + \varepsilon_{dbt}$ where $\overline{p}_{dbt}$ is the median price of a house in council d, band b at time t; τ_{dbt} is the average council tax for a house in council d, band b at time t; $\overline{Size}_{dbt}$ is the median size for a house in council d, band b at time t; δ_{bt} are band-time fixed effects; and $\overline{x}_{dbt}$ are council-band-time controls. Column (1) does not control for any variable; column (2) controls for the average size; column (3) controls for average size and adds number of rooms fixed effects; column (4) controls for average size, number of rooms and lighting, heating and hot water costs.

	(1)	(2)	(3)	(4)
Council Tax	-316.61***	-334.80***	-259.69***	-258.36***
	(18.42)	(18.54)	(62.07)	(62.19)
Size		1018.43***	1785.16***	1460.89***
		(153.82)	(471.28)	(521.21)
Lighting Cost				113.88
				(650.28)
Heating Cost				145.47***
				(46.94)
Hot Water Cost				-369.42
				(275.34)
FE	Year×Band	Year× Band	Year× Band	Year× Band
			Nb. Rooms	Nb. Rooms
R^2	0.76	0.77	0.83	0.83
Nb. Obs.	4,170	4,170	3,840	3,840

$4\% \times 316.61 = \pounds 12.66$. It is obvious that this figure is only the artefact of the negative correlation between the value of properties and the average tax in the council as observed in Figure 7. Extremely negative coefficients are obtained in columns (2), (3) and (4) where we control for the average size, number of rooms and energy costs, respectively. The smallest of these coefficients in absolute value, i.e., -258.36 in column (4), would imply an incidence of $r \times \beta = 4\% \times 258.36 = 10.33$ which is still unreasonably high given that, once properly discounted, the incidence should not be larger than one. The results in Figure 7 and Table 1 imply that any estimate of incidence needs to carefully take into account this spurious negative correlation: for this reason our identification strategy will try to compare only those houses that lie extremely close, i.e., no more than 300m and mainly closer than 100m, to the border to disentangle the actual incidence of the tax from the geographical distribution of taxes across councils. For the rest of the paper, the reader should bear in mind that the geographical distribution of council taxes entails that any estimated incidence will be, at most, an upper bound for the *true* incidence. This is because, if buyers tend to value certain characteristics upon purchasing a house, these should be capitalized in the house price which, in this case, acts almost like a sufficient statistic for their value; the results of Figure 7 and Table 1 signal that houses with better characteristics (and higher prices) tend to be located in councils with lower taxes, thus inflating any estimate of tax incidence. A second and more subtle reason why we can only estimate an upper bound for the incidence has to do with our identification strategy. By comparing similar dwellings on opposite sides of a border, we are implicitly assuming that the buyer always has an outside option during the price bargaining process. As a result, the buyer would be much more elastic than an otherwise identical buyer involved in the purchase of a house located in the heart of a council where there is no outside option in terms of council tax. We will show in Section 5 that the seller will bear the full incidence of the tax at the border, while that won't necessarily be the case at any other point. In general, even in the absence of perfect substitutes across council borders, it is reasonable to conjecture that the incidence will still be much larger at the border compared to the council centre, where the agent would have to move long distance in order to pay a different council tax rate.

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

3.2 Identification Strategies

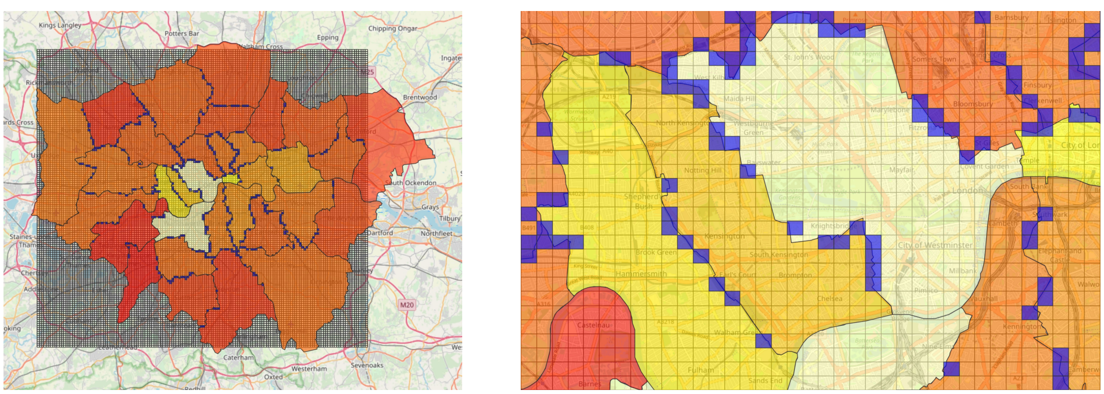
As mentioned previously, in this secton we are going to describe the two identification strategies that we are going to use to measure an upper bound of the incidence of council tax on property prices.

3.2.1 Regression Grids

The first strategy compares houses that lie in close proximity by dividing the area of London in a grid and assigning a fixed effect to each square in the grid. By doing so, we are de-facto comparing two houses that are otherwise identical but lie on opposite sides of a given border. Figure 8 graphically depicts our first approach. Panel 8a shows a map of London superposed by a grid of squares with equal size. Panel 8b shows a more detailed picture of the boroughs in the centre[9]. We then proceed to select the squares that have two houses that: are sold in the same year, are in the same council tax band and lie on

Figure 8 : Grids

The maps depict our startegy of dividing London in a grid of equally sized squares. Panel 8a shows a grid of 150×150 squares superposed to the map of the city; Panel 8b shows an enlargement of the central boroughs.



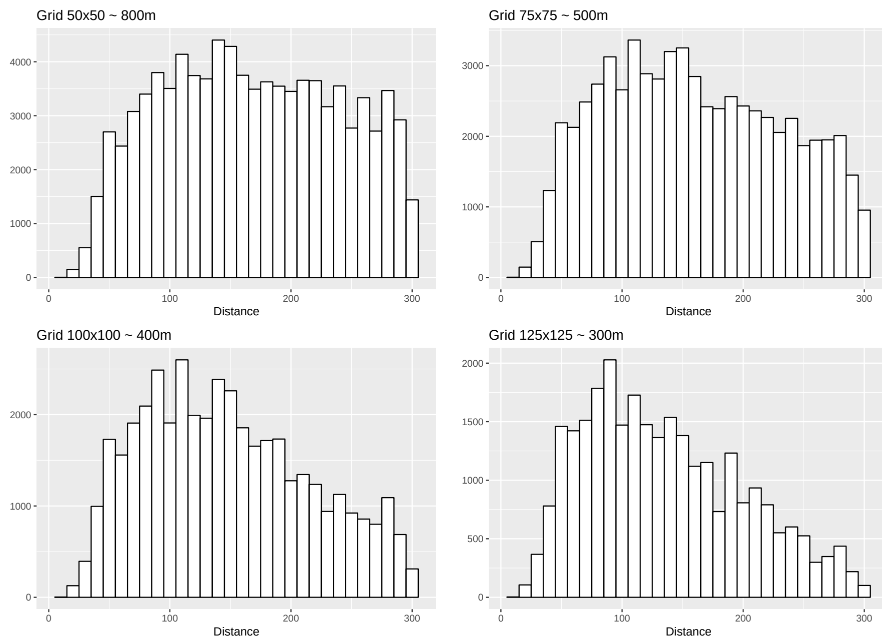
(a) : Grid **(b)** : Enlargement of the Centre

[9]The three main boroughs depicted in the picture are, starting from left, Hammersmith and Fulham, Kensington and Chelsea and Westminster.

opposite sides of the border; Panel 8b displays in blue examples of such squares. It can be noticed that we discard observations for which the border is located on the Thames River bank. To avoid relying on an arbitrary division, we have used four different grids,

Figure 9 : Distribution of Distances

The figure depicts histograms for the distribution of distances between houses that are used in our regressions. We report the distributions for four different grids, namely one where we have divided London in 50×50 squares, then one in 75×75, one in 100×100 and, finally, 125×125. For each histogram we report the approximate size of the square in meters.



.

namely one grid divides the area in 50×50 squares, another divides it in 75×75 squares, another in 100×100 squares and, finally, the last grid is a 125×125 one. These squares have an approximate size of 800 meters, 500 meters, 400 meters and 300 meters, respectively. While these are the maximal possible distances between houses, we have decided to remove observations that were more than 300 meters far from the border. Figure 9 shows the distribution of distances from the border for our different specifications. As mentioned, no house lies more than 300 meters away from the border, and most of the observations are about 100 meters away from the closest border. As we proceed to refine our grids by subdividing in a larger number of squares, we can see that we lose observations in the 100 meters-300 meters range; this will reduce our power significantly, but it will ensure that we compare houses that are indeed in very close proximity.

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

$$p_{ibgdt} = \alpha + \beta\tau_{bdt} + \gamma Size_i + \delta_{bgt} + \zeta' x_i + \varepsilon_{ibgdt} \quad (2)$$

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

3.2.2 Matching Estimator

Our second identification approach consists of pairwise matching of houses on opposite sides of a given border. To find the closest match, we need to define a distance: in what follows, we rely on a Euclidean distance and a distance based on a linear model. Under the first one, we restrict the possible matches to be within 500 meters distance from each other, to be sold in the same year, to be in the same council tax band, and to both be either old or newly built. For each property we then choose the closest match as the one minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$, where i is the original property, j indexes the possible matches on the other side of the border, x_{ik} are house i characteristics and x_{jk} are house j characteristics. We then run within-pair regressions:

$$p_{it} = \alpha + \beta\tau_{it} + \delta_{ijt} + \zeta' x_{it} + \varepsilon_{it} \tag{3}$$

where δ_{ijt} are ij-pair dummies, x_i are house i-specific features (size, number of rooms, energy cost).

The second choice of distance is based on a linear pricing model:

$$p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \tag{4}$$

where x_{it} contains the same house-specific features: size, number of habitable rooms and energy cost. We then compute the model-predicted price $\hat{p}_{it} = \hat{\alpha} + \hat{\beta}' x_{it}$. As before, we restrict the pairing to houses sold in the same year, band, old/new category and no further than 500m from each other. For each property we pick the closest match as the one that minimises the following distance: $d(i,j) = \hat{p}_{it} - \hat{p}_{jt}$. To estimate the incidence, we run within pair-regressions as in (3) where the δ_{ijt} dummies are determined according to the new matching algorithm. As in Section 3.2.1 the identification will be valid as long as the only systematic difference within pairs is the amount of council tax. As previously explained, any other omitted variable will most likely lead us to estimate an upper bound for the incidence, given the geographical distribution of council taxes.

4 Results

4.1 Grid Estimator

Table 2 presents the results of the grid regressions decribed in Section 3.2.1. In all specifications the coefficient on council taxes is economically very small and statistically insignificant. The first column compares the price of two properties in the same band that

Table 2 : Grid Regressions

.

	(1)	(2)	(3)	(4)
Council Tax	12.72	-0.42	-4.83	-0.51
	(8.50)	(8.58)	(8.52)	(8.40)
Size	4031.56***	4033.21***	3900.32***	4162.13***
	(39.80)	(39.76)	(40.21)	(59.09)
Lighting Cost				-1418.36***
				(58.25)
Heating Cost				171.89***
				(5.40)
Hot Water Cost				-168.27***
				(22.32)
FE	Grid × Year × Band	Grid × Year × Band	Grid × Year × Band	Grid × Year × Band
		Month	Month	Month
			Property Type	Property Type
R^2	0.72	0.72	0.75	0.76
Nb. Obs.	88,044	88,044	88,044	88,044

were sold in the same year on the opposite sides of a border, controlling for size measured in square meters. Notice that the council tax coefficient is insignificantly different from zero and it even has the wrong sign, implying that properties that pay higher taxes tend to also have higher prices. The coefficient on size informs us that the marginal price of a square meter is around $\pounds 4,000$. To control for seasonality, the second column adds month fixed effects to the previous specification. The coefficient on council tax remains statistically insignificant and, although negative, the point estimate implies an extremely low incidence: with an in-sample discount rate of about 4%, $r \times \beta = 0.04 \times (-0.42) = -0.0168$, i.e., for a $\pounds 1$ increase in the council tax paid per year the house price decreases by less than $\pounds 0.02$. This is considerably lower than the incidence of 1 that we should expect from rational agents living on the border. As a caveat, note that this exercise is a joint statement about tax incidence and discount rates used by agents. Conversely, to achieve an incidence of 1, one would have to believe that the average discount rate used by agents is 238% per year[10].In Section 5.1, we will use the estimates of the incidence of the stamp duty tax given by Best and Kleven (2018) in order to calibrate a model consisent with the above estimates and show that they will imply abnormally high discount rates. The third column also adds property type fixed effects, and the fourth adds additional controls for energy efficiency which may proxy for other unobservable characteristics of house quality such as maintenance. Even the largest estimate in column (3) would still imply discount rates higher than 20%. One may be tempted to attribute the magnitude and the statistical insignificance of the coefficient to lack of power. This cannot be the case given the very high R-squared, ranging between 72% and 76%, and the very precise estimates obtained for all other coefficients; for example, characteristics like lighting, heating and hot water cost exhibit very low standard errors and therefore large t-statistics, despite being similar in magnitude to council taxes. Unfortunately, frequentist statistics is ill-suited for proving that a coefficient is indeed equal to zero and it would be wrong to think that the confidence intervals provide us with a range for the true incidence given that they only tell us about $p(\hat{\beta}|\beta = 0)$. We are working on extending our analysis to a

[10]Notice that this is just a rough approximation of the true discount rate and it applies only if the agent discounts the future council tax payments as a perpetuity. The countercyclical nature of council taxes as shown in Figure 2 implies that it is very unlikely that these command a positive risk premium. Therefore, as we will assume in Section 5.1, the agents may discount future cash flows at a higher rate only because of the presence of liquidity constraints, but this will complicate the conversion between our point estimates and the implied discount rates.

Bayesian setting where we will be able to say more about $p(\beta|\hat{\beta}) \propto p(\beta) \times p(\hat{\beta}|\beta)$.

Table 3 : Grid Regressions - different grids

	(1)	(2)	(3)	(4)
Council Tax	-0.51	4.14	5.34	13.33
	(8.40)	(13.63)	(17.11)	(20.83)
Size	4031.56***	4059.43***	3657.38***	3323.09***
	(39.80)	(63.23)	(117.91)	(63.29)
Lighting Cost	-1418.36***	-1536.10***	-1358.44***	-1185.67***
	(58.25)	(61.01)	(125.00)	(148.01)
Heating Cost	171.89***	170.35***	166.41***	162.46***
	(5.40)	(5.72)	(10.06)	(12.65)
Hot Water Cost	-168.27***	-146.59***	-134.23***	-140.05***
	(22.32)	(23.55)	(44.43)	(52.74)
FE	Grid×Year× Band	Grid×Year× Band	Grid×Year× Band	Grid×Year× Band
	Month	Month	Month	Month
	Property Type	Property Type	Property Type	Property Type
R^2	0.76	0.70	0.77	0.77
Nb. Obs.	88,044	38,933	25,363	16,813
Grid	50×50	75×75	100×100	125×125

Table 3 displays the grid regression results for grids of different size: column (1) uses a grid that divides the London area into 50×50 squares, column (2) 75×75, column (3) 100×100 and column (4) 125×125. This might help to alleviate concerns that grids made of large squares might be comparing houses that are rather distant from each other. The specification is otherwise same as the one in column 4 of Table 2. The coefficient on council tax remains statistically insignificant and in most specifications it has the wrong sign. The fact that the R-squared is very high (between 70% and 77%) and that all other coefficients are precisely estimated confirms our previous finding that the incidence of the council tax is likely to be very small.

Table 4 : Grid Regressions - without stamp duty notches

	(1)	(2)	(3)
Council Tax	-1.47	3.58	1.66
	(8.54)	(8.79)	(8.72)
Size	4458.13***	4229.48***	4480.80***
	(65.49)	(60.62)	(65.99)
Lighting Cost	-1629.84***	-1476.71***	-1664.41***
	(64.19)	(59.74)	(64.73)
Heating Cost	184.40***	175.15***	185.50***
	(5.99)	(5.54)	(6.05)
Hot Water Cost	-166.09***	-166.22***	-164.09***
	(25.25)	(23.00)	(25.48)
FE	Grid×Year× Band	Grid×Year× Band	Grid×Year× Band
	Month	Month	Month
	Property Type	Property Type	Property Type
R^2	0.73	0.73	0.73
Nb. Obs.	78,216	85,627	77,766
$p \notin$	[240k-270k]	[490k-520k]	[240k-270k] & [490k-520k]

To make sure that the confounding effect of the stamp duty notches does not play a role in our estimation results, Table 4 presents the results of the grid regressions when we remove the two main stamp duty notches at $£250,000$ and $£500,000$. Column (1) excludes only the first notch, column (2) the second, and column (3) removes both. The results are virtually unchanged, with the incidence still being statistically insignificant, small in magnitude, and often displaying the wrong sign. As afore-mentioned, the large R-squared and the fact that the remaining coefficients are precisely estimated guarantees that this is not due to lack of power.

In the following section we are going to augment the evidence by presenting results using our second identification strategy.

4.2 Matching Estimator

Table 5 : Matching Regressions

	(1)	(2)	(3)	(4)
Council Tax	-7.41	-7.45	-8.65	-8.44
	(6.41)	(6.19)	(6.23)	(6.10)
Size	2053.74***	2025.65***	1937.06***	1850.20***
	(379.63)	(371.62)	(360.23)	(341.81)
Energy Cost	-499.98***	-698.52***	-481.95***	-701.95***
	(83.86)	(112.22)	(86.31)	(92.16)
Model	Euclidean	Euclidean	Linear	Linear
FE	Pair, Rooms	Pair, Rooms	Pair, Rooms	Pair, Rooms
Observations	122,328	122,328	122,328	122,328
R^2	0.838	0.839	0.841	0.841

Table 5 shows the results of our second estimation approach where we explicitly match similar dwellings on opposite sides of a border as described in Section 3.2.2. The first two columns display the results obtained using the Euclidean distance and the remaining two use the linear pricing model. The presence of pair fixed effects amounts to regressing the difference in prices of matched houses on the difference in council tax paid, controlling for size and energy cost as well as adding number of rooms fixed effects. None of the coefficients on council tax are statistically significant, although they are only slightly larger in magnitude relative to before. Even so, the largest estimate obtained in column (3) would still imply implausibly high discount rates. We are going to shed more light on these and the previous results in Section 5.1.

The empirical findings above demonstrate that council tax differences never significantly explain house price differences. While absence of evidence, namely the fact that agents seem to be insensitive to taxes that are postponed in the future, does not directly imply evidence of absence, there is plenty of other corroborating evidence: first, many point estimates are positive and hence with the wrong sign; second, the smallest coefficient obtained with the grids and matching approach is -4.83 and -8.65, respectively, implying a rather large discount rate. This is inconsistent with our calibrated model and would imply a time value of money larger than 18% (10%) in the worst case scenario and larger than 20% (12%) when we calibrate as in Section 5.1.

Bearing these estimates in mind, in the next section we develop a simple model that will allow us to propose a plausible explanation for the above results.

5 Model

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

$$\max_{\{c_t\}_{t=0}^{\infty}, h} \quad U(\{c_t\}_{t=0}^{\infty}, h) = c_0 + \sum_{t=1}^{\infty} \beta^t \left(u(c_t)\right) + \sum_{t=0}^{\infty} \beta^t \left(\frac{h^{1-\phi}}{1-\phi}\right) \tag{5}$$

$$s.t. \quad c_0 + h(p_{A0}\mathbb{1}_{\{A\}} + p_{B0}\mathbb{1}_{\{B\}} + \tau_S) \leq w_0 + d_0 \tag{6}$$

$$c_t + h(\tau_A \mathbb{1}_{\{A\}} + \tau_B \mathbb{1}_{\{B\}}) + d_{t-1}(1+r) \leq w_t + d_t \quad t = 1, 2, 3, ... \tag{7}$$

$$d_t \leq \alpha h(p_{At}\mathbb{1}_{\{A\}} + p_{Bt}\mathbb{1}_{\{B\}}) \quad t = 0, 1, 2... \tag{8}$$

For simplicity, the utility of the agent is chosen to be time-separable and separable in consumption and housing. The utility function is quasi-linear in c_0 in order to get rid

of income effects, as it is standard in the public finance literature. For tractability and to separate the effects of stamp duty and council tax, the agent purchases the housing good only once at $t = 0$. There are two councils, A and B, with exogenously chosen and potentially different council tax rates. We assume that there is equal supply of housing in both councils. Equation (6) is the first-period budget constraint: the agent spends his initial endowment w_0 on consumption c_0 and the after-tax cost of his housing demand h. When he buys a house, the agent pays the pre-tax price $p_{i0}, i = A, B$, and, in addition, he also needs to pay the stamp duty tax τ_S hereby assumed to be proportional to the quality-adjusted level of housing demand. If his total demand exceeds his initial endowment, the agent can borrow additional funds d_0 for one period at the risk-free rate. The budget constraints for all subsequent periods are identical and given by equation (7): from time $t = 1$ onwards, the agent will need to spend his endowment w_t on his optimal consumption choice c_t and to pay the council tax that corresponds to the council where he chose to locate $\tau_i, i = A, B$. He also needs to repay his short-term debt from the previous period with interest $d_{t-1}(1 + r)$, and he can borrow again at the same terms in order to balance his budget constraint. Finally, the last constraint in equation (8) is the financing constraint: the agent cannot borrow more than a fraction α of the pre-tax cost of his housing demand. This can potentially generate very large incidence for the stamp duty tax since the lump sum nature of this tax will tighten the leverage constraint.
The Lagrangian for the above problem can be written as:

$$\mathcal{L} = U(\{c_t\}_{t=0}^{\infty}, h) - \lambda_0(c_0 + h(p_{B0} + \tau_S) - w_0 - d_0) - \sum_{t=1}^{\infty} \lambda_t(c_t + h\tau_{Bt} + d_{t-1}(1+r) - w_t - d_t) - \sum_{t=1}^{\infty} \mu_t(d_t - \alpha h p_{Bt}) - h\mathbb{1}_{\{A\}} \left[\lambda_0(p_{A0} - p_{B0}) + \sum_{t=1}^{\infty} \lambda_t(\tau_{At} - \tau_{Bt}) + \alpha \sum_{t=0}^{\infty} \mu_t(p_{At} - p_{Bt}) \right] \tag{9}$$

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

$$p_{A0} - p_{B0} \le -\sum_{t=1}^{\infty} \frac{\lambda_t}{\lambda_0}(\tau_{At} - \tau_{Bt}) + \alpha \sum_{t=0}^{\infty} \frac{\mu_t}{\lambda_0}(p_{At} - p_{Bt}) \tag{10}$$

i.e., if the price differential between the same-quality house in councils A and B more than compensates for the present value of the difference in future council tax payments and the collateral value of the house. In equilibrium, markets clear if equation (10) holds with equality which, from now onwards, we assume to be the case. Assuming that the agent is indifferent between living in councils A and B, we proceed by suppressing the council subscripts and denote the price of the house as p and the council tax as τ. The first-order conditions for an interior solution are:

$$1 = \lambda_0 \tag{11}$$

$$\beta^t u'(c_t) = \lambda_t \quad \forall t = 1, 2, 3, ... \tag{12}$$

$$-\lambda_t + \lambda_{t+1}(1+r) + \mu_t = 0 \quad \forall t = 0, 1, 2, ... \tag{13}$$

$$\frac{h^{-\phi}}{(1-\beta)} = \lambda_0(p_0 - \alpha\frac{\mu_0}{\lambda_0}p_0 + \tau_S) + \sum_{t=0}^{\infty}\lambda_{t+1}\tau_{t+1} - \sum_{t=0}^{\infty}\lambda_{t+2}\frac{\mu_{t+1}}{\lambda_{t+2}}\alpha p_{t+1} \tag{14}$$

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

$$\frac{\lambda_{t+1}}{\lambda_t} = \beta\frac{u'(c_{t+1})}{u'(c_t)} = \frac{1}{1 + r + \frac{\mu_t}{\lambda_{t+1}}} \tag{15}$$

The above Euler equation implies that the agent's discount factor is equal to the inverse of the risk-free rate and a liquidity premium $\frac{\mu_t}{\lambda_{t+1}}$, arising from the fact that the house has some collateral value. In order to simplify the exposition, we assume that in equilibrium the liquidity premium is constant and equal to $\frac{\mu_t}{\lambda_{t+1}} = k$, and that house prices grow at a constant rate g, i.e., $p_t = p_0(1+g)^t$. Re-arranging equations (10), (14) and (15), we obtain the final no-arbitrage condition and housing demand:

$$(p_{A0} - p_{B0})\left(1 - \frac{\alpha k}{r+k-g}\right) = -(\tau_{A0} - \tau_{B0})\frac{1}{r+k} \tag{16}$$

$$\frac{h^{-\phi}}{(1-\beta)} = p_0\left(1 - \frac{\alpha k}{r+k-g}\right) + \tau_S + \frac{\tau}{r+k} \tag{17}$$

The first equation is the equilibrium condition of how house prices should behave across councils, i.e., the house price differential, after having taken into account the collateral value $\frac{\alpha k}{r+k-g}$, needs to match (the negative of) the present value of the council tax differential. The second equation states that the agent's marginal utility of housing is equal to the house price inclusive of (the present value of) all taxes and collateral value. It is important to note that the no-arbitrage condition (16) could, in principle, give a different incidence compared to the one obtained from the housing demand (17). This is because the former holds only at the border between two councils where the outside option, i.e., the option of buying an otherwise identical house on the other side of the border, implies that the supply bears the whole burden of the tax. In particular, we obtain an incidence of:

$$\frac{dp}{d\tau} = -\frac{1}{r+k} \times \frac{r+k-g}{r+(1-\alpha)k-g} \tag{18}$$

On the other hand, for both houses on the border as well as houses in the middle of a given council we can define the optimal demand from equation (17) as $D(p, \tau, \tau_S) = h^*(p, \tau, \tau_S)$. Equating with the optimal supply, $S(p) = D(p, \tau, \tau_S)$, and after total differentiation gives us the standard formula for the incidence:

$$\frac{dp}{d\tau} = -\frac{\frac{\partial D}{\partial \tau}}{\frac{\partial D}{\partial p} - \frac{\partial S}{\partial p}} = -\frac{1}{r+k} \times \frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \tag{19}$$

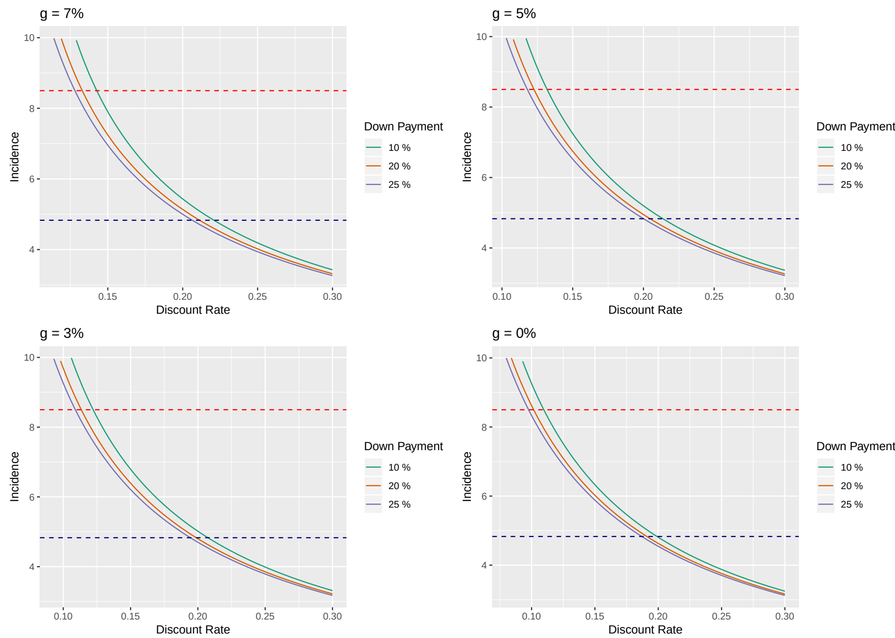
where $\tilde{\eta}_S = \frac{\partial S}{\partial p}\frac{p}{S}\frac{p\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau}{r+k}}{p}\phi = \eta_S\frac{p\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau}{r+k}}{p}\phi$ is a slightly modified version of the supply elasticity η_S that takes into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

$$\frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \leq \frac{r+k-g}{r+(1-\alpha)k-g} \tag{20}$$

implying that the incidence at the council border is an upper bound for the *true* council tax incidence as long as the modified elasticity of supply is non-negative, i.e., $\tilde{\eta}_S \geq 0$. Notice that the modified elasticity of supply $\tilde{\eta}_S$ is positive as long as the true elasticity of supply η_S is positive and the difference between the two increases as the structural elasticity of demand $\frac{1}{\phi}$ decreases in magnitude.

5.1 Calibration

Figure 10 : Calibrated Incidence



The model in the previous section allows us to better interpret the empirical results of Section 4. Equation 18 suggests that to calibrate the implied discount rate we need to measure the liquidity premium k, the down payment $1-\alpha$ and the growth rate of house prices g. The liquidity premium k can be calibrated from the results in Best and Kleven (2018) by noticing that the model-implied incidence of the stamp duty tax is[11]:

$$\frac{dp}{d\tau_S} = -\frac{r+k-g}{r+(1-\alpha)k-g} \tag{21}$$

Figure 10 shows the implied risk-free rate r as a function of the estimated incidence β. The different panels plot this relationship for various housing price growth rates, namely for $g = 7\%, 5\%, 3\%, 0\%$. Note than the average growth rate in sample is 7.2% with the median rate being equal to 7.3%, implying that our choices are rather conservative. For each panel we depict the relationship for three different values of the down payment parameter $(1-\alpha) = 10\%, 20\%, 25\%$ and these are consistent with typical values in the UK mortgage market. The horizontal lines represent the largest estimates for the incidence:

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

the blue one is equal to the largest estimate obtained from the grid regressions (-4.83), while the red one is equal to the largest estimated derived from the matching approach (-8.65). Looking at the first panel one can see that the largest coefficient obtained from the grid regressions implies a risk-free rate of more than 20% per year. The matching algorithm implies a risk-free rate of more than 12% per year. It should be stressed that these are the most conservative estimates for the discount rate which, in general, will be larger when we use the other point estimates. It can also be noticed that the risk-free rate is not very sensitive to different growth rate assumptions, i.e., moving from the first to the last panel still implies very large discount rates, with the lowest one being equal to 10%.

The results above become striking once coupled with the extent to which house buyers react to stamp duty notches. The evidence that the incidence of the stamp duty tax is large compared to the incidence of the council tax could, in principle, be due to two things: a large liquidity premium k or a large risk-free rate r. The fact that the incidence of the stamp duty is large but not extreme implies that the liquidity premium cannot be the only source of the low council tax incidence. This leads us to believe that, when buying their properties, agents discount tax payments that happen in the future disproportionately compared to tax payments that occur concurrently with the purchase of the property. It is difficult to argue that this might be due to risk associated with council tax payments given their countercyclical nature, as can be seen in Figure 2. This leaves us with the last possible option: agents fail to take fully into account the council tax upon purchasing a property, either because this is much less salient compared to the stamp duty tax[12], or because they fail to appreciate the magnitude of its present value[13]. Notice also that the results so far suggest that there is somebody who is not taking the council tax differentials into account in a fully-rational way, but this does not need to be the house buyer: our previous analysis goes through even if the buyer is fully aware of the tax and hopes to shift its incidence onto the subsequent buyer, or the renter in the case of buy-to-let property transactions.

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

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

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

5.2 Implications for Tax Policy

Given the results in the previous section, it seems reasonable to argue that agents fail to fully perceive deferred taxes unless one is willing to assume implausibly high discount rates. As a result, we propose a modified version of the model above that allows for non-fully salient taxes. For simplicity, let us assume that the growth in housing prices equals zero, i.e., $g = 0$. Let us also assume we are in an equilibrium where the leverage constraint (8) is binding, i.e., $d_t = d = h\alpha p$. If we multiply each of the constraints (6) and (7) by $\frac{1}{(1+r+k)^t}$ and add them together, we obtain the following consolidated budget constraint:

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \tag{22}$$

where $\tilde{p} = p\left(1 - \frac{\alpha k}{r+k}\right) + \tau_S + \frac{\tau}{r+k}$ is the tax-inclusive house price. Following Chetty et al. (2009), Farhi and Gabaix (2015) and Goldin (2015), we assume that the agent misperceives taxes with attenuation factor m, i.e., he solves the following maximization problem:

$$\max U(\{c_t\}_{t=0}^{\infty}, h) = c_0 + \frac{h^{-\phi}}{1-\phi} + \sum_{t=1}^{\infty} \beta^t \left(u(c_t) + \frac{h^{-\phi}}{1-\phi}\right) \tag{23}$$

s.t.

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}_m h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \tag{24}$$

where the perceived house price is:

$$\tilde{p}_m = p\left(1 - \frac{\alpha k}{r+k}\right) + \tau_S + m\frac{\tau}{r+k} \quad m \in [0,1] \tag{25}$$

Notice that while the agent perceives the above budget constraint, he has to satisfy the actual budget constraint (22) given by the rational model. As pointed out in Reck (2016), it is crucial to decide what choice variable will bear the burden of adjustment. Given our assumption about the quasi-linear utility function in first-period consumption c_0, it is natural to let c_0 be the shock absorber. This choice amounts to assuming the following train of events: 1) the agent misperceives the council tax he will have to pay going forward and, as a result, buys ”too much” quality-adjusted housing; 2) following this, he realizes that the actual amount of taxes he will have to pay is beyond his budget; 3) consequently, the agent adjusts his consumption in the first period keeping everything else constant. Denoting the observed demands as $\hat{c}_0, \hat{c}_t, \hat{h}$, and the optimal demands absent any behavioural frictions as c_0^*, c_t^*, h^*, we have the following first-order conditions:

$$\hat{c}_t = [u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right) = c_t^* \tag{26}$$

$$\hat{h} = [(1-\beta)\tilde{p}_m]^{-\frac{1}{\phi}} \neq [(1-\beta)\tilde{p}]^{-\frac{1}{\phi}} = h^* \tag{27}$$

$$\hat{c}_0 = I - \sum_{t=1}^{\infty} \frac{\hat{c}_t}{(1+r+k)^t} - \hat{h}\tilde{p} \neq c_0^* \tag{28}$$

As previously mentioned, the optimality condition for future consumption remains as before. However, equation (27) shows that the agent will demand too much housing due to the fact that the perceived price $\tilde{p}_m$ is lower than the true price $\tilde{p}$, as long as $m < 1$. As a result, because of quasi-linearity in the utility function, $\hat{c}_0$ will adjust to absorb the reduction in available income. The previous discussion highlights the fact that misperception of the house price will affect both consumption and housing demand, albeit in opposite direction. This implies that a benevolent social planner needs to carefully balance the two distortions when setting the optimal tax policy. To see this more formally, let us adopt the approach of Goldin (2015) and assume that the government will choose the optimal (property) tax combination in order to raise a fixed amount of revenue and maximize the utility of the buyer. For convenience, define the present value of council tax revenue from the Government's point of view, discounted at the risk-free rate, as $\tilde{\tau} = \frac{\tau}{r}$. The total revenue raised from a given buyer is:

$$R = (\tau_S + \tilde{\tau})h \tag{29}$$

The Government can twick the two taxes to maintain revenue-neutrality. In particular, a revenue-neutral tax change will be such that:

$$\left[h + \tau_S \frac{\partial h}{\partial \tau_S}\right] \Delta\tau_S = -\left[h + \tilde{\tau}\frac{\partial h}{\partial \tilde{\tau}}\right] \Delta\tilde{\tau} \tag{30}$$

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

$$\frac{\Delta\tau_S}{\Delta\tilde{\tau}} = -\frac{h + \tilde{\tau}\frac{\partial h}{\partial \tilde{\tau}}}{h + \tau_S\frac{\partial h}{\partial \tau_S}} = -\frac{h + \tilde{\tau}\theta_{\tilde{\tau}}\frac{\partial h}{\partial p}}{h + \tau_S\theta_{\tau_S}\frac{\partial h}{\partial p}} \tag{31}$$

where $\theta_{\tau_S} = \frac{\frac{\partial h}{\partial \tau_S}}{\frac{\partial h}{\partial p}}$ and $\theta_{\tilde{\tau}} = \frac{\frac{\partial h}{\partial \tilde{\tau}}}{\frac{\partial h}{\partial p}}$ tell us how responsive the demand is with respect to taxes relative to pre-tax prices. Next, define $k_1 = (1 - \frac{\alpha k}{r+k})$ and $k_2 = \frac{r}{r+k}$. The indirect utility function for an inattentive agent will be:

$$V(p, \tau_S, \tilde{\tau}, r, k) = I - \sum_{t=1}^{\infty} \frac{\hat{c}_t}{(1+r+k)^t} - \hat{h}(k_1 p + \tau_S + k_2\tilde{\tau}) + \sum_{t=1}^{\infty} \beta^t u(\hat{c}_t) + \frac{\hat{h}^{(1-\phi)}}{(1-\beta)(1-\phi)} \tag{32}$$

where $\hat{c}_t = \hat{c}_t(r,k) = [u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right)$ and $\hat{h} = \hat{h}(p, \tau_S, \tilde{\tau}) = [(1-\beta)(k_1 p + \tau_S + m k_2 \tilde{\tau})]^{-\frac{1}{\phi}}$ from the agent's first-order conditions. Differentiate the indirect utility function above to obtain:

$$\frac{dV}{d\tilde{\tau}} = -\hat{h}\left(k_1\frac{\partial p}{\partial \tilde{\tau}} + k_2 + \frac{\partial \tau_S}{\partial \tilde{\tau}}\right) + \left[\frac{\partial U}{\partial h} - (k_1 p + \tau_S + k_2\tilde{\tau})\right]\left[\theta_{\tilde{\tau}} + \frac{\partial p}{\partial \tilde{\tau}} + \theta_{\tau_S}\frac{\partial \tau_S}{\partial \tilde{\tau}}\right]\frac{\partial \hat{h}}{\partial p} \tag{33}$$

As in Goldin (2015), the change in welfare can be decomposed into four components: the first part, i.e., $-\hat{h}\left(k_1\frac{\partial p}{\partial \tilde{\tau}} + k_2 + \frac{\partial \tau_S}{\partial \tilde{\tau}}\right)$ measures the direct welfare effect of a tax shift due to the alleviation of the borrowing constraint; the second part, i.e., $\left[\frac{\partial U}{\partial h} - (k_1 p + \tau_S + k_2\tilde{\tau})\right]$ is the behavioural wedge and it represents the difference between perceived and actual prices; the third component, i.e., $\left[\theta_{\tilde{\tau}} + \frac{\partial p}{\partial \tilde{\tau}} + \theta_{\tau_S}\frac{\partial \tau_S}{\partial \tilde{\tau}}\right]$ is equal to the change in prices as perceived by the agent; and the fourth component, i.e., $\frac{\partial \hat{h}}{\partial p}$ is the impact of a change in prices on demand for housing. With no bias, i.e., when $m = 1$ the perceived price is equal to the actual price and the envelope theorem ensures that the second component above is equal to zero. As a consequence, given that the first term is always positive, it is optimal for the government to set $\tau_S = 0$. This is because by doing so, the Government helps alleviate the agent's liquidity constraint. In the presence of biases, however, there is a trade-off between the two inefficiencies: 1) the liquidity constraint and differences in salience make raising the stamp duty tax less efficient that raising the council tax; 2) on the other hand, raising the council tax causes a shift in demand away from c_0 which in our example is the shock absorber. The problem of the social planner, therefore, will be to choose the optimal combination of stamp duty and council tax to jointly solve the following two equations:

$$\hat{h}(\tau_S + \tilde{\tau}) = R \tag{34}$$

$$\frac{dV}{d\tilde{\tau}} = 0 \tag{35}$$

In practice, changing the tax mix, i.e., the combination of council tax and stamp duty tax rates, can change the inattention parameter m.

6 Conclusions

This paper studies the incidence of property taxes in the UK housing market. By using a geographical discontinuity approach, exploiting the considerable difference in council tax rates across London Boroughs, we show that economic agents significantly under-react to council taxes. Given the empirically observed typical loan-to-value ratio and growth rate of house prices, the risk-free rate implied by our results is above 10% in the most conservative scenario, and above 20% in more realistic scenarios. This is in sharp contrast to the large stamp duty incidence estimated by Best and Kleven (2018) and suggests that agents do not pay sufficient attention to taxes deferred to the future, or possibly points to evidence of very large search frictions or other cognitive costs. In Section 5.2, we have briefly touched upon the policy implications of our findings, however, one should be aware of issues arising when manipulating tax rates given that there is no guarantee that changes in policies are not followed by changes in tax salience and therefore behaviour. The analysis in this paper relies on data from the residential property market, however, it can also be extended to other domains of tax policy. One general take-away from the present work is that transaction taxes, such as the stamp duty tax, have a large incidence on transaction prices while deferred taxes, such as the council tax, have a lower effect on prices but potentially higher impact on consumption choices. This implies that the optimal mix of taxes may be some combination of the two. The analysis can be extended, for instance, to financial securities where the fact that a transaction tax might be very distortionary does not imply that it is optimal to raise revenues only through capital gains[14] or dividend taxes.

[14] While the capital gains tax is a transaction tax, the fact that it is borne by the seller of the asset suggests that agents could still underreact to it as it is a deferred tax and, therefore less salient compared to a tax

In future versions of this paper, we aim to tackle remaining issues; first of all, the analysis in Section 4 needs to be expanded to properly test, in a Bayesian setting, whether it is likely that the true incidence is zero; second, we are planning to explore differences in consumption responses at the border which we should expect to arise whenever agents fail to optimally account for tax differences and are forced to adjust their expenditures ex-post to meet their budget constraints.

charged at the moment of purchase like the stamp duty tax.

References

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

Besley, Timothy, Neil Meads, and Paolo Surico, "The incidence of transaction taxes: Evidence from a stamp duty holiday," *Journal of Public Economics*, 2014, *119* (C), 61–70.

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

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

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

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," NBER Working Papers 21524, National Bureau of Economic Research, Inc September 2015.

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

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

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

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

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

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

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