Evasión y Agrupamiento en el Impuesto Territorial: Evidencia de Chile

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA

Alumna: Catalina María Gómez Colomer
Profesor Guía: Eduardo Engel Goetz

Santiago, diciembre de 2021

Preliminary version, please do not cite

Abstract

This is the first paper to estimate and model property tax evasion in Chile. Property size as reported to Chile's national tax authority by homeowners is compared with estimates of property size obtained from a private appraisal company in order to produce estimates of underreporting. These estimates are rationalized via an optimal tax evasion model where a tax notch at 140 square meters (after which the tax rate on properties doubles) plays a central role. The model parameters are estimated using method of moments. Moments that capture bunching for both actual and reported square meters at the tax notch are consequential for the estimation. Finally, the model is used to analyze the implications of various changes to the property tax rate, particularly moving to a regime with no tax notch. A regime with no tax notch results in a significant reduction in tax evasion.[1]

In 2019, Chile's national broadcaster aired an investigation exposing the prevalence of property tax evasion among the country's high-income groups, including serving politicians. The broadcast, using administrative and anecdotal evidence, described the different ways in which individuals manage to evade property tax, including not registering the property with Chile's national tax authority (known by its Spanish acronym, SII), declaring a false use of the property (i.e., agricultural rather than residential), or underreporting the area of buildings by not regularizing new extensions to the property. Property tax evasion is particularly important because it is a progressive tax; it consists of three increasing tax rates for three brackets of property value. Moreover, property tax is the main source of income for municipalities, amounting to nearly 38% of municipalities' income, particularly for lower-income administrative areas[2].

One Chilean law known as DFL-2 (1959) aims to support social and affordable housing by reducing by half the rate of tax on properties of 140 square meters or less[3]. This law creates a tax notch—that is, a discontinuity in the average tax rate—and as a result it creates an incentive to underreport the size of buildings. This poses the question, how does the tax notch defined by the DFL-2 impact the behavior of households regarding the real and reported size of their properties?

In this paper the focus is on evasion that is achieved by underreporting the size of the property. An optimal tax evasion model is proposed that incorporates a discontinuity in the tax rate with close form solutions. A novel database allows the measurement of evasion by size of the property through access to the official SII data that records reported size and new data from a private appraisal company that provides an estimation of the actual sizes of properties.

[1] This thesis is part of a Master's degree in Economics at the Universidad de Chile. Supervising Professor: Eduardo Engel.

[2] Tax revenues are distributed from higher-income residents to lower-income residents, and property tax correlates with wealth

[3] The tax rate doubles for properties of more than 140 square meters. This also applies to inframarginal square meters. The corresponding tax rate for properties depends on an appraisal. Properties with appraisals that are equal to or less than US$41,000 are exempt of the property tax. For properties with appraisals equal to or less than US$146,000, the annual tax rate is 0.933%, and for property appraisals over US$146,000 the annual rate is 1.088%.

This is the first study in Chile that measures and analyzes tax evasion of property tax. The findings show evasion of significant magnitude (7.6% of the real size) and various effects of the DFL-2 law on evasion. More specifically, households bunch near the threshold and this behavioral response is seen particularly in reported size suggesting an important evasion response because of the DFL-2. This prompted the development of a model of property tax evasion, where a rational individual chooses the size of their property and how much to underreport to the authority. An individual pays a rate equal to t_1 if the reported size of the property is more than u and pays a rate equal to $t_0 < t_1$ if the property is less than u. With a probability $p(e) = p_0 e$, an audit will occur discovering the full amount of evasion and the individual must pay the amount evaded plus a proportional fine.

The collected data is used to estimate the model allowing a study of the effects of three counterfactual scenarios with a constant rate, revoking the tax notch. The tax rates evaluated are t_0, t_1, and a tax rate that maintains aggregate utility unchanged denoted by $\bar{t}$, where $t_0 < \bar{t} < t_1$. In all cases evasion decreases significantly and tax revenue increases when the rate is equal to t_1 and $\bar{t}$. This suggests a number of potential policies that will help to reduce evasion and increase the tax income of municipalities.

There are five main contributions to the existing literature produced by this study. First, it is the first study of property tax evasion using Chilean data. Second, it provides evidence of the effects of a property tax law that provokes a tax rate discontinuity—that is, bunching near a notch for real and reported property size. Third, a model of tax evasion with a tax notch, and with close form solutions, is developed. Fourth, the model is estimated with the data obtaining a good fit for the main characteristics. And fifth, the model is used to analyze counterfactual scenarios where the tax notch does not exist, which decreases the evasion rate and in some cases increases tax revenues.

The paper is organized as follows: Section 1 presents the related literature. Section 2 explains property tax in Chile. Section 3 introduces the data along with four stylized facts. Section 4 develops an static model of tax evasion with the tax notch, which is later estimated by a method of moments in Section 5. Finally, Section 6 present the counterfactual exercises and 7 concludes.

1 Relation to literature

This paper primarily relates to a recent body of literature that studies the bunching phenomenon. Bunching occurs in the presence of a discontinuity in the tax rate and can occur as a result of a tax kink—a discrete change in the marginal tax rate—or a tax notch—a discontinuity in the average tax rate. Recent literature has used bunching for a new empirical approach to estimate structural parameters (Kleven, 2016). This type of analysis, based on kink points, was developed by Saez (2010) who examined kink points in the US income tax schedule and built a model to estimate compensated elasticities using bunching evidence. Kleven & Waseem (2013) developed a design based on notches.

Bunching responses can occur through legal or evasion responses, or both, which is very relevant for this paper. Saez (2010) finds clear evidence of bunching around a kink point for self-employment income but no evidence of bunching for dependent workers that only receive income from their wages. This suggests that most of the bunching may be because of reported rather than real income effects. Seim (2017) studies the behavioral response to wealth taxes in Sweden, which includes financial securities, real estate, and consumption durables. The paper examines the response to a kink point from self-reported and third-party-reported wealth finding smaller bunching for individuals with no incentive to misreport (third-party and taxable wealth below the exempt threshold), which suggests substantial underreporting to evade the tax. Using income data from Denmark, Kleven et al. (2010) study a tax enforcement field experiment where a selection of individuals were audited. This allows a break down of the total response to a kink into evasion and legal response. In the pre-audit distribution, Kleven et al. (2010) observe substantial bunching around the kink, bigger than the observed bunching in the post-audit distribution, which implies there is an evasion response as well as a real response to the kink. This work also extends the standard economic model of tax evasion so that the probability of auditing is dependent on the type of income being underreported (third-party versus self-reported). The model predicts low evasion for third-party income and larger effects of tax enforcement for self-reported income. Literature regarding property tax and bunching effects is less prevalent. Albouy et al. (2021) focus on a kink produced by an exemption base on the size of property in the rental market in Tehran. Using this kink the price elasticities of housing supply and demand are estimated using administrative data and a structural model.

The theoretical literature follows the main models of tax evasion. It begins with the work of Allingham & Sandmo (1972), who develop a static model where the actual income is exogenously given and the tax payer decides the amount they will declare to the authorities. However, as a result of a certain probability an audit will occur, the taxpayer will have to pay the undeclared amount at a penalty rate higher than the tax rate.

Finally, with respect to the literature on property tax in Chile, this is the first study that examines and estimates evasion. Furthermore, it is the first study that focuses on the behavioral response of tax notches in the tax system in Chile and, specifically, the tax notch introduced by the DFL-2.

2 Property Tax in Chile

The property tax is a tax that must be paid on real estate based on the value of the real estate. The administration and auditing is the responsibility of the SII; however, the revenue from property tax goes to municipalities. The base of this tax is a downward biased estimate of the property's market value, which is a function of the area, location, and quality of material of the property, among other factors. The property tax is progressive and very relevant for the finances of Chilean municipalities. In the case of nonagricultural real estate for residential use, the tax rate increases according to three brackets of appraisal:properties that are appraised as being equal to or less than US$41,000 are exempt from property tax, for properties that are appraised as being equal to or less than US$146,000 the annual tax rate is 0.933%, and properties over US$146,000 the annual rate is 1.088%. Over 60% of properties in Chile are exempt of property tax as a result of being appraised as under US$41,000.

Property tax in Chile correlates with wealth: the richest municipalities in the Metropolitan Region have more than 95% of residential properties paying taxes, whereas in the lowest-income areas less than 7.6% of households pay property tax. Also, the correlation between the proportion of households that do not pay property tax and level of poverty in the municipalities in the Metropolitan Region is positive and high (0.68). In addition, municipalities have a system of horizontal transfers contributing 60% of the revenues from property tax to a common fund, the five richest municipalities contribute 65%. The fund is distributed according to the total income of the municipality and the proportion of residents that live below the poverty line. Thanks to this mechanism property tax is the main source of income for municipalities, it's nearly 38% of municipalities' income, and a substantial part of the common fund equating to almost 60% in 2017 (Razmilic et al., 2015; Department of Treasury of Chile). Residential properties correspond to approximately two thirds of the total taxable properties and one third of the property tax revenues, this accounts for the difference in valuation relative to other types of properties (e.g., commercial) and various exemptions (Razmilic et al., 2015).

In the last five years, a number of news articles and television programs in Chile have revealed how high-income individuals, including politicians, have been evading the property tax. Evasion can be accomplished by different means, as described by anecdotal evidence. First, some individuals may not register their property with the authorities and will therefore not be taxed. Second, an individual could register a false use for the property. For example, by declaring a property to have an agricultural use instead of a residential one, since agricultural usage is subject to lower taxes. Another method, common according to the anecdotal evidence, is to underreport the area of the buildings on a property. This can be achieved by not reporting new extensions to the authorities or hiding rooms using fake walls that are later demolished. The penalties for underreporting are stipulated in the General Law of Urbanism and Construction that establishes fines for buildings without regulation and in the Tax Code that stipulates penalties for evasion. When there is an incomplete or erroneous declaration that leads to less taxes being paid the fine is between 5% and 20% of the resulting difference.

One particular law introduces an incentive to underreport size of the buildings on a property. The Decree-Law N°2 (DFL-2, 1959) provides tax relief for residential properties that are 140 square meters or less. This decree aimed to support social and affordable housing. Properties that meet this criteria have a 50% exemption from property tax[4] for up to 20 years, which is the focus of this work. Other benefits of this law include an exemption from taxes when selling, renting, donating, or inheriting the property[5]. Accordingly, the DFL-2 introduces a *tax*

[4]The SII will apply the most advantageous exemption for the tax payer. There is a limit of two properties per person. And the exemption expires after 20 years if the property is 70 built square meters or less, 15 years if it is between 70 and 100, and 10 years if it is between 100 and 140.

[5]This only applies to properties with residential purposes and with natural persons as owners. With respect to inheriting and donating, this exemption can be used only if the donor has built or acquired the property in the first transfer.

notch—that is, a discontinuity in the average tax rate. Given the anecdotal evidence and the DFL-2, this study will focus on the underreporting of the area of the buildings as a relevant means of evasion.

3 Data

Our main objective is to estimate the magnitude of evasion and its behavior given the DFL-2. Particularly, we need a proxy for underreported square meters (sqm). Data from the SII have the information of all register properties in the country (for several years) with their main characteristics such as size, age and taxes paid. The novelty of this work is that we have access to a data that supply an estimate of actual sqm. The data come from a private appraisal company that estimate the value of property. In the process it obtains fundamental information like size, constituting an independent measure from SII registers. Each property have a unique identification number that allow us to merge both data sets. As we know the exact date of valuation from the private appraisal, we can compare with the SII data in the same semester. We will denote the sqm from SII as "reported sqm", and the sqm from the private company as "real sqm".

The first best would have been to measure the taxes paid with the taxes that the household would actually have to pay if they share truthful information to the authority, meaning, if we have a counterfactual for the fiscal appraisal. Unfortunately, we don't have all the information that would allow us to calculate the counterfactual fiscal appraisal with the data from the private company, since this appraisal depends not only on square meters but also on price of the land defined by the SII, quality of materials, among others. However, as explained before, underreport in the size of the property is not only a viable practice to evaded the property tax, but is relevant considering the tax notch introduced by the DFL-2.

3.1 Basic Descriptive Statistics

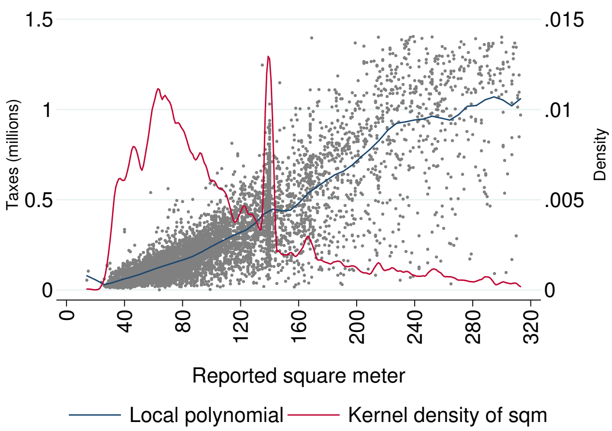
In order to work with a more homogeneous sample, that at the same time has enough properties subject to property tax, we focus on a subsample (N=8,241): residential properties from Metropolitan Region municipalities with more than 400 observations and more than 70% of taxable properties. The data is from the years 2015 (50%), 2016 (28%) and 2017 (22%) and considers eight municipalities: Las Condes, Santiago, Ñuñoa, Lo Barnechea, Providencia, Vitacura, San Miguel and La Reina[6] which represent 24% of the total population of the Metropolitan Region and earn more than 37% of the property tax revenues of the whole country (National Statistics Institute of Chile; Department of Treasury of Chile).

Because there isn't a one-to-one correspondence between sqm and taxes paid, we begin by describing their relation. Figure 1 shows a positive relationship summarize in the local polynomial, although it has a great variance. It is important to notice that there are properties of all sizes, including those with reported sqm below 140, that pay taxes suggesting that small properties could also have

[6]Proportion in the sample respectively: 27%, 18%, 13%, 11%, 11%, 10%, 5% and 4%

appraisals over the exempt value. Besides, the figure shows a kernel density of the reported sqm, which described the distribution of properties where the great majority relies on and below 140 meters.

Figure 1. Semiannual taxes against reported sqm



Note: the blue line is a Kernel-weighted local polynomial smoothing of degree 2 and bandwidth 5, the kernel function use is epanechnikov. The red curve is a kernel density with bandwidth 2.

Figure 2 provides evidence of underreporting which seems relevant for its frequency and magnitude. First, the underreporting happens at all levels of reported square meterage. Second, the average underreporting is 7.6% of real sqm and more than half of homeowners underreport more than 7%. Nevertheless, there is a considerable fraction that underreport close to zero: 19% have underreport between -1 and 1. Finally, some individuals have reported sqm greater than its real sqm which means negative underreporting (9% has underpreport under -1). This phenomenon could occur because of measurement error or difference in the measuring method between SII and private appraisal company. Finally, Table 1 presents the main descriptive statistics for our central variables: not surprisingly, reported sqm have a smaller mean than real sqm.

Figure 2. Underreport sqm against reported sqm

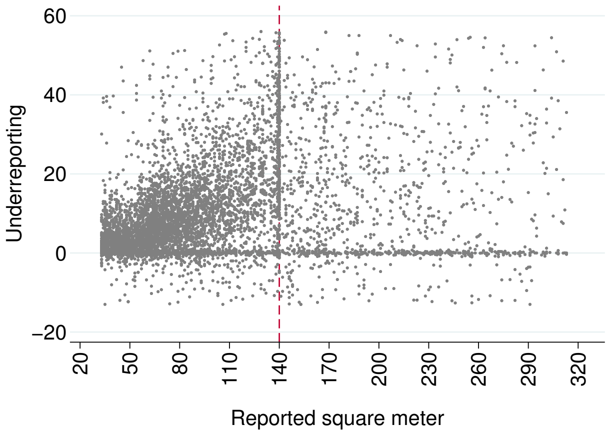


Table 1. Descriptive statistics: real, reported and underreported sqm

	Real	Reported	Underreported
Mean	131.0	121.0	10.0
Standard Deviation	108.6	110.2	50.8
Median	99.3	90.0	6.0
Interquartile range	100.0	80.0	15.4

3.2 Stylized Facts

This section establish four key facts showed in our novel data.

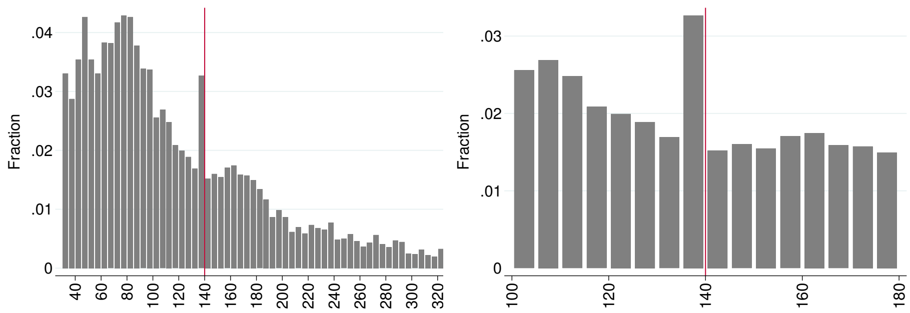
1. Households bunch at the DFL-2 threshold for real sqm

DFL-2 introduce a big incentive to have a property of 140 square meter or less to save half of the property taxes. Figure 3 presents a histogram of the real sqm where we can see bunching of households around 140. The fraction of households that choose to be at 140 more than doubles the households that choose to be right above 140. In fact, 1.6% have properties from 139 to 140 whereas only 0.2% from 140 to 141. Measurement error, difference in sizing methods and (or) frictions of the real world, could make the spike smaller than it would be if no error or frictions exists[7].

[7]Notice that near 70 or 100 sqm, the other two thresholds for DFL-2, don't exhibit such notorious bunching; this serve as a confirmation that the salient and relevant notch for individuals is the one at 140. This could be because the threshold 140 defines the property tax exemption as well other tax benefits, whereas at 70 and 100 sqm only changes the duration of the property tax benefit.

Figure 3. Histogram for real sqm

(a) From 35 to 320 sqm **(b)** From 100 to 185 sqm



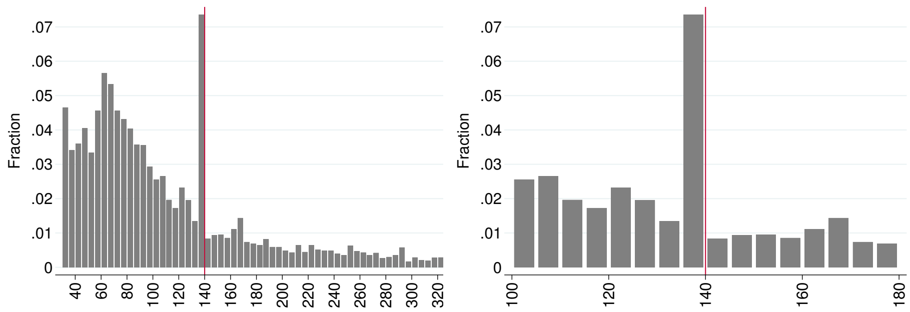
Note: Bin width = 5. Domain: $(30, 330]$

2. Bunching at the threshold is larger for declared sqm

In Figure 3 we show the histogram of reported sqm. Just as in real sqm, there is bunching around 140 because there is also an incentive to *report* having a property of 140 sqm or less. The bunching is more than twice as large as the one found in real sqm. Therefore, DFL-2 seems to introduce a larger incentive of reporting rather than actually having a property of 140 sqm or less. The behavioral response of taxpayers is distributed between real and reporting responses. Fact N° 1 and 2 are the most important for what follows.

Figure 4. Histogram for reported sqm

(a) From 35 to 320 sqm **(b)** From 100 to 185 sqm

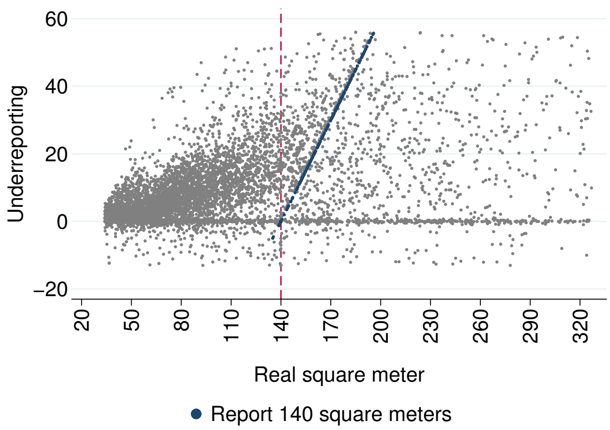


Note: Bin width = 5. Domain: $(30, 330]$

3. Underreport increases sharply after the threshold, as a function of real sqm

Households that choose to report 140 square meters have very different levels of real sqm. In Figure 5 underreporting increase after 140 as a function of real sqm, we highlight those with reported sqm equal to 140.

Figure 5. Underreported sqm against real sqm

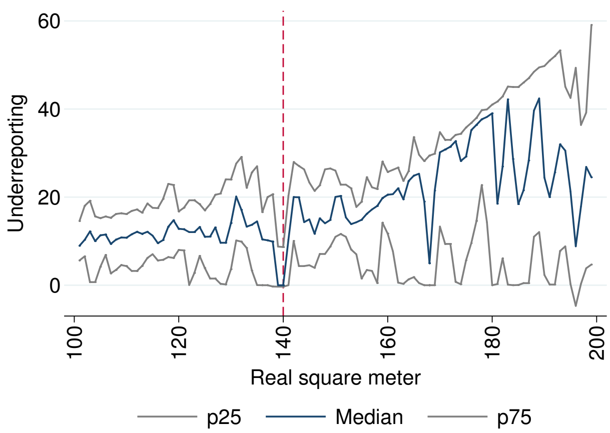


Note: The blue dots are properties that report exactly 140 sqm.

4. Underreport drops near zero around the threshold, as a function of real sqm

Underreport as a function of real sqm falls close to zero at 140 sqm as shown by Figure 6. One possible explanation is that individuals who have a optimal property near 140 decide to underreport just enough to report 140 so they can benefit from DFL-2 and minimize the costs of underreporting that is greater for those who have real sqm over 140, but reported sqm below 140, as evasion law establishes that fines are proportional to the amount evaded.

Figure 6. Median of underreport against real sqm



Note: p25 and p75 denotes the 25th and 75th percentiles respectively. The percentiles were calculated taking overlapping intervals of width two.

4 A model of property tax evasion

In what follows we present a model of evasion where households simultaneously choose the size of their property and how much to underreport to the authority. The goal is to replicate the stylized facts describe in Section 3.2, mainly the bunching phenomenon.

4.1 Setting and case with no notch

A rational individual works and produces $\sqrt{A}L$ units, where A is her productivity and L are the hours work. Her disutility from work is $\frac{L^2}{2}$. She only consume one good: her property. The utility is defined by $w(C) = C$ where w is the utility function and C is the size of her property. We denote by z the sqm of the property. Assuming preferences are additively separable and since all production goes into the household $z = \sqrt{A}L$, then without taxes the household maximizes:

$$g(z) = z - \frac{z^2}{2A}, \tag{1}$$

Property owners pay a tax with rate t_0 on reported sqm. Denoting the later by d and underreporting by e, so that $z = d + e$, taxes pay are $T(z - e)$ with $T(d) = t_0 d$. With probability $p(e) = p_0 e$ ($p_0 > 0$) an audit occurs discovering the full magnitude of evasion. If audited the household must pay the amount evaded $T(z) - T(z - e)$ plus a fine proportional to it with parameter $\theta \in (0, 1)$. The structure of the tax fines is motivated by the tax laws.

It follows that the property owner maximizes over z and e:

$$U(z, e) = z - \frac{z^2}{2A} - T(z - e) - p(e)(1 + \theta)[T(z) - T(z - e)] \tag{2}$$

So the solutions are:

$$\hat{z} = (1 - t_0)A \tag{3}$$

$$\hat{e} = \frac{1}{2p_0(1 + \theta)} \tag{4}$$

4.2 Introducing a Tax Notch

Now we introduce what it is called in the literature as "tax notch". This is a discontinuity in the tax schedule $T(z)$, where the average tax rate change after a threshold (u). Unlike a "tax kink" where there is a increase in the marginal tax rate (Kleven, 2016). The new tax schedule is:

$$T(d) = \begin{cases} t_0 d & \text{if } d \leq u \\ t_1 d & \text{if } d > u \end{cases} \tag{5}$$

With $t_0 < t_1$. In our case $u = 140$ and $t_0 = t_1/2$. Solving (2) subject to (5) for z and e we obtain close form solutions:

Proposition 1. *(Optimal z)*

$$z^* = \begin{cases} (1-t_0)A & \text{if } A < A_0 \\ u & \text{if } A \in [A_0, A_1) \\ \frac{A[1+up_0(1+\theta)(t_1+t_0)]}{1+2p_0 t_1 A(1+\theta)} & \text{if } A \in [A_1, A_2) \\ (1-t_1)A & \text{if } A > A_2 \end{cases} \quad (6)$$

Proposition 2. *(Optimal e)*

$$e^* = \begin{cases} \frac{1}{2p_0(1+\theta)} & \text{if } A < A_1 \\ z-u & \text{if } A \in [A_1, A_2) \\ \frac{1}{2p_0(1+\theta)} & \text{if } A > A_2 \end{cases} \quad (7)$$

Proposition 3. *(Optimal d)*

$$d^* = \begin{cases} (1-t_0)A - \frac{1}{2p_0(1+\theta)} & \text{if } A < A_1 \\ u & \text{if } A \in [A_1, A_2) \\ (1-t_1)A - \frac{1}{2p_0(1+\theta)} & \text{if } A > A_2 \end{cases} \quad (8)$$

We have explicit although complicated formulas for each threshold of A: A_0, A_1 and A_2 that depends on parameters θ, t_1, t_0, p_0, u. See Appendix 8.1 for the proof of the solutions, the expressions for A_0, A_1 and A_2, and more details.

4.3 Model implications

We identify the consequences of the tax notch by showing the relationship between z, d and e with respect to the productivity A in the case with and without tax discontinuity (Figure 7). The real and reported size of the house (z and d) are increasing functions of the productivity A, but not strictly as there are ranges of A for which z and d are constant. The discontinuous increase in the average tax rate produce a range of productivities where the household prefers to maintain z (or d) constant equal to the threshold u. This implies: less work because of the smaller z chose for the same productivity; less taxes paid because of the smaller reported size d chose for the same productivity.

On the other hand, e is constant[8] except for those with $d^* = u$, where is an increasing function of A. These individuals choose $z^* > u$ and $d^* = u$, so they are subject to the tax rate t_0 instead of t_1 which is the rate accordingly to their real size z. For this reason, in case of an audit they would have to pay an amount $(1+\theta)[T(z) - T(d)] = (1+\theta)[t_1 z - t_0 d] > (1+\theta)t_0 e$. The marginal cost of e is larger for those with $d = u$ and $z = u + \varepsilon$ for a small $\varepsilon > 0$, than for those with $d = u$ and $z = u$. As a result, the individual would want to minimize the cost of underreport and still be subject to the smaller rate t_0.

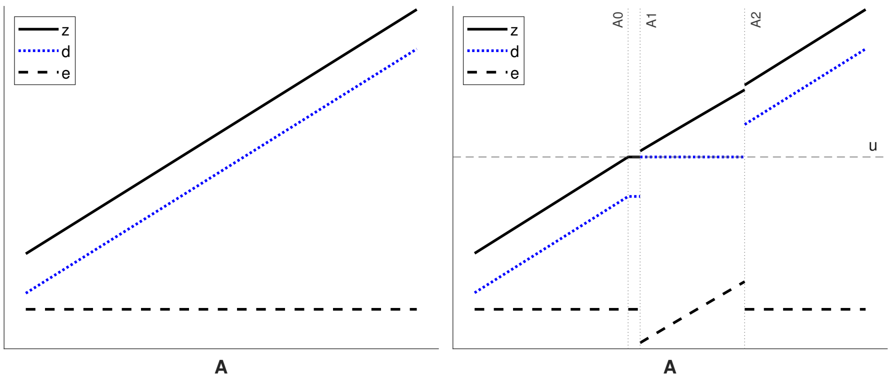
From these implications we can understand the aggregate implications of our model. The aggregate version of our model consist on assuming that A is a

[8] Another probability function that depends negatively on z like $p(e,z) = p_0 \frac{e}{z^\gamma}$ with $0 < \gamma < 1$ lead to a different result with e being increasing in A. We didn't do this to obtain close form solutions. The main conclusions doesn't change.

random variable that is distributed according to a continuous distribution function such as a lognormal with parameters μ_A and σ_A, we assume that these parameters are equal for all individuals and the variance is large enough to have individuals affected by the tax notch. The main aggregate implication is that households bunch at the threshold u. The bunching occurs both in z and d, this implies the existences of density hole just above u. Refer to Figures 8 and 9 to see the distribution of d and z for some given parameters. There is a also density hole for d below u. Under certain conditions[9] d jumps from its optimal value when $z^* = u$, to the value $d^* = u$, the reason behind this behavior is the same previously explained: marginal cost of evasion is larger when the individual choose $d^* \leq u$ and $z^* > u$.

Figure 7. Optimal choices: z, d, e against A

(a) Without tax notch **(b)** With tax notch



5 Model Estimation

We are able to set in advanced the following parameters:

- $t_1 = 0.933\%$, according to the SII.
- $t_0 = 0.4665\%$, because the DFL-2 stipulates that properties pays only half (in this case $t_0 = t_1/2$).
- $u = 140$, according to the DFL-2.
- $\theta = 0.125$, $\theta \in [0.05, 0.2]$ according to the fines specifies in tax laws.

Nevertheless we don't have a clear measure for p_0 or A. The parameter p_0 will be estimated. Besides, because z is an (not strictly) increasing function of A and due to the form of observed distribution of z (see Figure 3) we will assume A has a continuous distribution that takes positive values as the standard lognormal function, where μ_A and σ_A the mean and variance of the logarithmic values, respectively.

[9]See Appendix 8.1.

$$A \sim \text{lognormal}(\mu_A, \sigma_A)$$

5.1 Method of Moments

To estimate the model parameters we will use the method of moments. Denoting by N the total number of observations (equal to 8,241), from Fact N° 1 and 2, DFL-2 provokes bunching near the threshold, we have these moments:

$$\hat{m}_1 = \frac{\sum_{i=1}^{N} 1(z(i) \in [138, 140])}{N}$$
$$\hat{m}_2 = \frac{\sum_{i=1}^{N} 1(d(i) \in [135, 140])}{N}$$

To better assess the spread of the distribution of productivity (σ_A), we also include:

$$\hat{m}_3 = \frac{\sum_{i=1}^{N} 1(z(i) > 140)}{N}$$
$$\hat{m}_4 = \frac{\sum_{i=1}^{N} 1(d(i) > 140)}{N}$$

As noted in the basic descriptive statistics, there is a considerable fraction of home owners with small underreported amount. To account for this feature we extend the model to two types of individuals. A fraction $1 - \alpha$ as before, and fraction α choose z optimal and does not underreport. This introduces a new parameter α to estimate and motivates a fifth moment ($\hat{m}_5$):

$$\hat{m}_5 = \frac{\sum_{i=1}^{N} 1(e(i) \in (-1, 1))}{N}$$

In synthesis we have four parameters to estimate: μ_A, σ_A, p_0 and α; and five moments to match. We denote $\beta = (\mu_A, \sigma_A, p_0, \alpha)$. For every β we calculate $m_1(\beta)$, $m_2(\beta)$, $m_3(\beta)$, $m_4(\beta)$ and $m_5(\beta)$ and minimize the loss function called "average absolute deviation":

$$v(\beta) = \frac{\sum_{j=1}^{N} |\log(m_j(\beta)) - \log(\hat{m}_j)|}{N} \tag{9}$$

Where $m_j(\beta)$ is the estimated moment and $\hat{m}_j$ the data moment. Since we have explicit expressions for d and z as a function of A we are able to calculate explicit expressions for $\log(m_j(\beta))$ and no simulations are needed.

5.2 Results

The results of the estimation are shown in Table 2 and 3. We estimate two models: one where $\alpha = 0$ excluding m_5 ($e \in [-1, 1]$), the second where $\alpha \geq 0$, including m_5. For both estimations the average absolute deviation is small, all theoretical moments are close to their counterpart in the data. Also, the addition of α makes the calibration more realistic, since the model on its own it's not capable of reproducing the behavior of the underreport we see in Figure 2, thus we select the second model. According to Bernasconi (1998) the audit probability is normally between 0.01 and 0.03 which match our estimations.

Table 2. Parameters

	$\alpha = 0$	$\alpha \geq 0$
μ_A	4.62	4.67
	(0.023)	(0.020)
σ_A	0.72	0.59
	(0.044)	(0.034)
p_0	3.03%	2.94%
	(0.20%)	(0.20%)
α	0.00%	18.68%
	(-)	(0.43%)

Note: Bootstrap standard deviations in parenthesis.

Table 3. Data and estimated moments

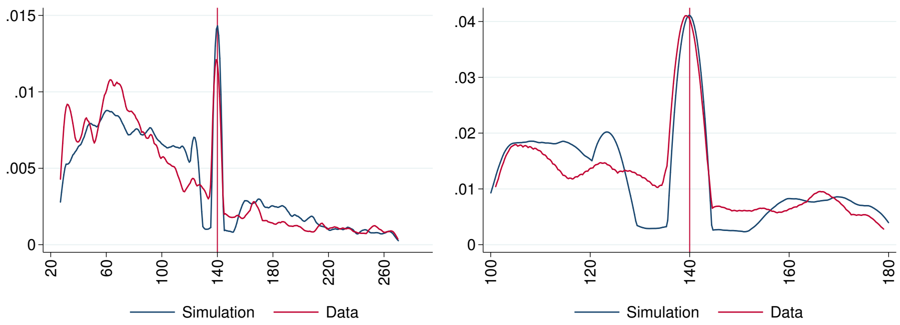
	Data	$\alpha = 0$	$\alpha \geq 0$
Average absolute deviation	–	0.0152	0.0125
m_1 ($z \in [138, 140]$)	2.03%	2.03%	2.03%
m_2 ($d \in [135, 140]$)	7.07%	7.07%	7.07%
m_3 ($z > 140$)	32.29%	30.39%	30.34%
m_4 ($d > 140$)	23.31%	23.31%	23.32%
m_5 ($e \in (-1, 1)$)	18.67%	0.00%	18.68%

5.3 Evaluation of Estimation

The smooth kernel estimation of the data and simulated distributions evidence how well our estimated model fit the data (see Figure 9 and 8). Stylized facts N° 1 and N° 2 are incorporated in the estimation, because they are the most important features we want to reproduce, which is accomplished as seen in the figures.

Figure 8. Kernel density of data and simulated model: d

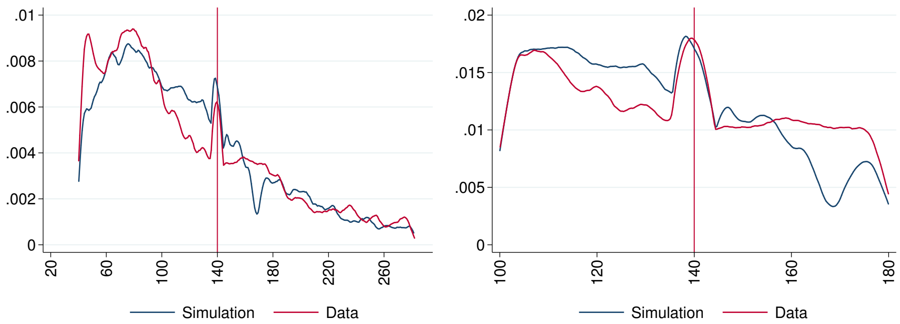
(a) **(b)**



Note: Epanechnikov kernel; half-width of kernel = 2.

Figure 9. Kernel density of data and simulated model: z

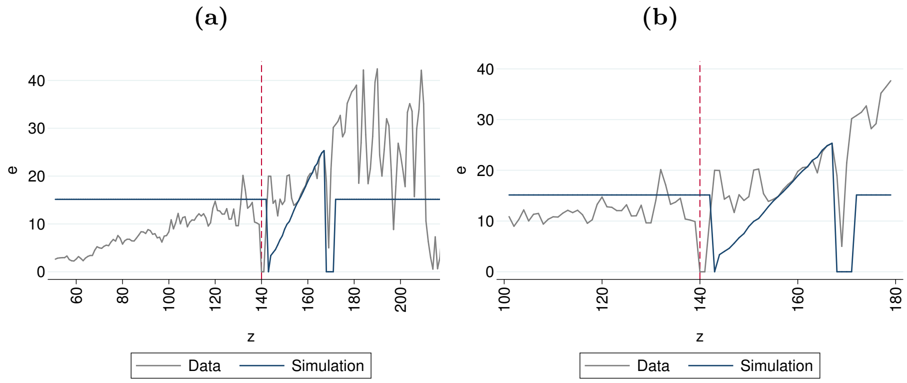
(a) **(b)**



Note: Epanechnikov kernel; half-width of kernel = 2.

For the evaluation of Fact N° 3 and N° 4 we compare the smooth median of e against z described in Figure 10. Far from the threshold the fit is not good because our model describe a constant underreport whereas the data show an increasing underreport for small values of z. Nevertheless our focus is near the threshold where our simulation does well for the most part. In particular the Figure 10 shows a sharp decrease around 140 followed by a rapid increase in underreport after the threshold.

Figure 10. Median of e against z around the threshold



Note: p25 and p75 denotes the 25th and 75th percentile. The smooth median and percentiles are calculated taking overlapping intervals of width two.

Given the basic descriptive statistics that shows negative underreporting and suggest a potential measurement error in the estimation of actual property size by the private appraisal company, we develop an estimation of the model only with the SII data obtaining good fit. See Apendix 8.2 for details.

6 Counterfactuals

In comparison to a tax schedule without a tax notch, a schedule with a tax notch can create inefficiencies because it imposes a very high marginal tax rate over a range of workers. There are three mechanisms through which the tax notch may lead to inefficiencies. First, less work because of households that choose a smaller z for the same productivity i.e those who bunch at the threshold; second, less taxes paid because there is also bunching near the threshold for d, these individuals choose a smaller d for the same productivity; third, less taxes paid as larger evasion occurs for individuals who choose $d \leq u$ and $z \geq u$ because these individuals pay tax at a rate equal to t_0 instead of their actual rate $t_1 > t_0$.

The implications of the tax notch led to an analysis of three counterfactual scenarios in which the tax notch no longer exists and, instead, all individuals are subject to the same tax rate, equal to: (i) the tax rate above the threshold $t_1 = 0.933\%$; (ii) the tax rate for those on or below the threshold $t_0 = 0.4665\%$; (iii) and a constant tax rate that maintains the aggregate utility constant denoted by $\bar{t} = 0.7025\%$. Note that $t_0 < \bar{t} < t_1$.

The individuals most affected by these changes are the ones who chose $d_0 = u$ with a tax notch because they reported a property size which pays t_0, when they should have been charged t_1. In consequence, they have a significant level of tax evasion. To illustrate the importance of these individuals, they represent 7% of all individuals but 46% of all tax evasion defined as $T(z) - T(d)$. At an aggregate level, the evasion rate ($\frac{\sum_{i=1}^{N}[T(z_i)-T(d_i)]}{\sum_{i=1}^{N} T(z_i)}$) is 13%. For the group with

$d_0 = 140$ it is 54%, this serves as a reference point for the importance of the tax notch for property tax evasion.

The most direct consequence of the tax notch ceasing to exist is that there is no bunching around the threshold. In the three counterfactuals, the decrease in evasion (Table 4) is primarily due to changes in the group that chose $d_0 = 140$ in the baseline model (Table 5). Regarding tax revenue ($T(d)$), naturally the larger increase occurs when the tax rate charged is t_1. Conversely, there is a decrease in tax revenues when the tax rate for all individuals is t_0. For a tax rate $\bar{t}$, tax revenue increases by 2.7%. When the analysis only takes account of individuals with d_0, tax revenue increases in all cases.

The impacts described are a lower bound because audits are costly. This factor is not accounted for in the model. A lower evasion rate could imply less audits and less resources expended on this task. For the same reason, we focus on tax revenue that does not account for fines because in an ideal world there would be no evasion, no audits, and no fines. In addition, evasion is costly for individuals because they have to make an effort to evade tax (e.g., hiding rooms, moral costs, etc.) which can also implies benefits of reducing evasion not accounted in the model.

These counterfactual exercises show that the revocation of the tax notch has a significant effect on tax evasion. For individuals that are affected by the notch ($d_0 = 140$) tax revenue increased for all and the evasion rate decreased considerably. At an aggregate level, the tax revenue increased for counterfactuals with tax rate t_1 and $\bar{t}$; however, it decreased for t_0 because there is a significant proportion of these individuals not affected by the notch who use to pay t_1 and now pay half the amount.

Table 4. Percentage change from baseline model to counterfactual scenarios

Tax rate	t_1	t_0	$\bar{t}$
Evasion rate	-27%	-27%	-28%
Tax Revenue	36.0%	-31.6%	2.7%
Utility	-0.4%	0.5%	0%

Note: Evasion rate is define as $\frac{\sum_{i=1}^{N}[T(z_i)-T(d_i)]}{\sum_{i=1}^{N}T(z_i)}$, *Tax Revenue as* $\sum_{i=1}^{N}T(d_i)$; $t_0 = 0.4665\%$, $t_1 = 0.933\%$, $\bar{t} = 0.7025\%$.

Table 5. Percentage change from baseline model to counterfactual scenarios: for $d_0 = 140$

Tax rate	t_1	t_0	$\bar{t}$
Evasion rate	-82%	-82%	-82%
Tax Revenue	103%	1.9%	53%
Utility	-0.4%	0.5%	0.04%

Note: Evasion rate is define as $\frac{\sum_{i=1}^{N}[T(z_i)-T(d_i)]}{\sum_{i=1}^{N} T(z_i)}$, *Tax Revenue as* $\sum_{i=1}^{N} T(d_i)$*;* $t_0 = 0.4665\%$, $t_1 = 0.933\%$, $\bar{t} = 0.7025\%$*; This table shows the changes for individuals with optimal reported size equal to* 140 *when there is a tax notch.*

7 Summary and Concluding Remarks

Property tax only represents a small proportion of the total tax revenue in Chile and is therefore often neglected. It is an important tax, however, because of its progressiveness and significant contribution to the income of Chilean municipalities. It has become evident that property evasion occurs, and it is particularly adopted by high-income groups. The Chilean taxation system provides an interesting context for this study because of the DFL-2 law that stipulates all properties of a size equal to or lower than 140 square meters receive a 50% exemption from property tax. This produces a discontinuity in the tax rate and an incentive to underreport size. This, in addition to anecdotal evidence, suggests that underreporting of size is a relevant means of property tax evasion.

This study employs a novel database that enables the measurement of underreporting of property size because it combines the reported square meterage to the Chilean tax authority and the information from a private appraisal company that supply estimations of the actual size of properties. Because of this combination of data an analysis of how real and reported size behave in light of the DFL-2 is possible. Evidence of underreporting (7.6% of real size) and of bunching near the threshold for real and reported size was found. The bunching for reported size is higher than for real size, suggesting that evasion plays a relevant role. Driven by these features, a model of tax evasion was developed introducing a tax notch with close form solutions, which was then estimated with a simulated method of moments using the novel database.

The primary policy implications stem from the counterfactual exercises that are applied using the simulated model. The revocation of the tax notch is evaluated by instead implementing a constant tax rate. Three rates are investigated: (i) the tax rate charged above the threshold t_1, (ii) the tax rate charged below the threshold t_0, and (iii) the tax rate where aggregate utility remains constant (denoted by $\bar{t}$ with $t_0 < \bar{t} < t_1$). The tax schedule without tax notch decreased the evasion rate significantly (by around 27%). Considering the importance of property tax and the evasion evidenced in this study, a revision of the DFL-2 would be beneficial. Originally, the purpose of the DFL-2 law was to support affordable housing; however, the market has changed and small properties can now have large market values. Besides, a lot of exemptions already exist that result in less than 40% of properties being subject to taxes. A reform of property tax law could have relevant effect. This study suggests that there is a policy that can decrease evasion and potentially increase tax revenues for municipalities, this will have a direct benefit of the lowest-income municipalities.

This is the first study of property tax evasion in Chile and the first to focus on the DFL-2 and its effects on real and reported size. The model developed in this paper does not reproduce all data characteristics and future research could develop a more complex model that replicates other interest features, such as variation in underreported size given a certain productivity. Extensions of the model could include the introduction of dynamic effects in the model as anecdotal data suggests that unregulated extensions are common—that is, a household can choose their property size in an initial stage and later decide by how much to increase the size of their property. On the other hand, richer data could enable the estimation of a counterfactual fiscal appraisal using the information from the private appraisal company to have a measure of tax evasion. Additionally, other means of evading property tax were not studied because of the lack of data. For instance, the false declaration of purpose (agricultural versus residential) seems to be an important source of evasion considering the difference in tax rates between the two uses and should be studied in the future. Finally, it may be worthwhile to investigate the role of municipalities in evasion because they directly benefit from decreasing property tax evasion but must also confront a potential political trade-off: a strong anti-evasion policy could hurt a mayor's reelection bid. This situation is amplified because it is often the case that those who should be subject to more tax are the same individuals who finance political campaigns. Until recently, mayors—the main authorities in municipalities—had no limit to the number of times they could be reelected. A new law in Chile has now limited the reelection of mayors, resulting in a maximum of three periods in office. This change could be exploited by evaluating the audit behavior of municipalities with mayors who can no longer be reelected with majors that can still run for reelection.

References

Albouy, D., Hurtado, C., & Nafari, K. (2021). Supply and demand responses to a tax on rental housing: Evidence from iran.

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Taxation: critical perspectives on the world economy*, *3*, 323–338.

Bernasconi, M. (1998). Tax evasion and orders of risk aversion. *Journal of Public Economics*, *67*(1), 123–134.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*, 435–464.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2010). Unwilling or unable to cheat. *Evidence from a Randomized Tax Audit Experiment in Denmark. pEconometrica,(forthcoming)*.

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

Razmilic, S., et al. (2015). Impuesto territorial y financiamiento municipal. *Estudios Públicos*, *138*, 47–91.

Saez, E. (2010). Do taxpayers bunch at kink points? *American economic Journal: economic policy*, *2*(3), 180–212.

Seim, D. (2017). Behavioral responses to wealth taxes: Evidence from sweden. *American Economic Journal: Economic Policy*, *9*(4), 395–421.

8 Appendix

8.1 Model solution

8.1.1 Minimization of G(e)

First, we fix $z > u$. Without loss of generality, $t_0 = 1$ and $t_1 = t > 1$. We define $G(e)$ by:

$$G(e) = \begin{cases} e^2 p_0(1+\theta)t + t(z-e) & \text{if } e < z-u \\ e^2 p_0(1+\theta) + e[p_0 z(1+\theta)(t-1) - 1] + z & \text{if } e \geq z-u \end{cases}$$

For now on:

$$G_1(e) = e^2 p_0(1+\theta)t + t(z-e)$$
$$G_2(e) = e^2 p_0(1+\theta) + e[p_0 z(1+\theta)(t-1) - 1] + z$$

$$G'(e) = 0 \rightarrow \begin{cases} e_1^*(z) = \frac{1}{2p_0(1+\theta)} & \text{if } e < z-u \\ e_2^*(z) = \frac{1-p_0 z(1+\theta)(t-1)}{2p_0(1+\theta)} & \text{if } e \geq z-u \end{cases}$$

$$z - u = e_1^*(z) \rightarrow z = \frac{1}{2p_0(1+\theta)} + u = z_1$$
$$z - u = e_2^*(z) \rightarrow z = \frac{1 + 2up_0(1+\theta)}{p_0(1+\theta)(t+1)} = z_2$$

We will assume:

$$u > \frac{1}{p_0(1+\theta)(t-1)}$$

This implies that $e_2^*(z) < 0$ so $z - u > 0 > e_2^*(z)$. For simplification we will focus in this particular case that doesn't lose the important features of our model.

Case I: $e_1^*(z) > z - u \leftrightarrow z < z_1$

$min[G_1(z-u), G_2(z-u)] = G_2(z-u)$

Because $G_1(z-u) > G_2(z-u) \leftrightarrow z < \frac{1+up_0(1+\theta)}{p_0(1+\theta)}$ condition that is met because $z < z_1 < \frac{1+up_0(1+\theta)}{p_0(1+\theta)}$

Case II: $e_1^*(z) < z - u \leftrightarrow z > z_1$

$min[G_1(e_1^*), G_2(z-u)]$

The result is a quadratic equation with one root that is greater or equal to u:

$$\bar{z} = \frac{t + p_0 u + p_0 t u + p_0 \theta u + p_0 t \theta u}{2 p_0 t(\theta + 1)}$$
$$+ \frac{p_0^{1/2} u^{1/2} [(t-1)(2t - p_0 u + p_0 t u - p_0 \theta u + p_0 t \theta u)]^{1/2} (\theta + 1)^{1/2}}{2 p_0 t(\theta + 1)} \geq u$$

So:

$$min(G(e(z))) = \begin{cases} G_2(z-u) & \text{if } z < \bar{z} \\ G_1(e_1^*) & \text{if } z > \bar{z} \end{cases}$$

Proposition 4 (Optimal evasion)**.** *Be* $v = (p_0, t, u, \theta)$*. Given the therholds* $\bar{z}(v)$*:*

$$e^*(z) = \begin{cases} z-u & \text{if } z \in (u, \bar{z}] \\ \frac{1}{2p_0(1+\theta)} & \text{if } z > \bar{z} \end{cases} \tag{10}$$

8.1.2 Maximization of H(z)

Given $e(z)$ from (10) and including the possibility for $z < u$ the individual maximizes:

$$U(z, e(z)) = H(z) = z - \frac{z^2}{2A} - G(e(z))$$

$$H(z) \begin{cases} H_1(z) = z - \frac{z^2}{2A} - t_0(z - \frac{1}{2p(1+\theta)}) - p_0(1+\theta)t_0(\frac{1}{2p(1+\theta)})^2 & \text{if } z \leq u \\ H_2(z) = z - \frac{z^2}{2A} - t_0 u - p_0(1+\theta)(z-u)[t_1 z - t_0 u] & \text{if } z \in (u, \bar{z}] \\ H_3(z) = z - \frac{z^2}{2A} - t_1(z - \frac{1}{2p(1+\theta)}) - p_0(1+\theta)t_1(\frac{1}{2p(1+\theta)})^2 & \text{if } z > \bar{z} \end{cases}$$

$$H'(z) \begin{cases} H_1'(z) = 1 - \frac{z}{A} - t_0 & \text{if } z \leq u \\ H_2'(z) = 1 - \frac{z}{A} - p_0(1+\theta)[t_1 z - t_0 u] - p_0(1+\theta)(z-u)t_1 & \text{if } z \in (u, \bar{z}] \\ H_3'(z) = 1 - \frac{z}{A} - t_1 & \text{if } z > \bar{z} \end{cases}$$

$$H'(z) = 0 \rightarrow \begin{cases} z_1^* = (1-t_0)A & \text{if } z \leq u \\ z_2^* = \frac{A[1+up(1+\theta)(t_1+t_0)]}{1+2p_0 t_1 A(1+\theta)} & \text{if } z \in (u, \bar{z}] \\ z_3^* = (1-t_1)A & \text{if } z > \bar{z} \end{cases}$$

$H(z)$ is not continuous in u, but it is continuous in $\bar{z}$ because $G_1(e_1^*(\bar{z})) = G_2(\bar{z} - u)$

First, $\frac{1}{2p(1+\theta)}$ is different from $\frac{1-pz(1+\theta)(t-1)}{2p(1+\theta)}$, except when $t = 1$. So $H(z)$ is discontinue in u.

Second, $\frac{1-pz_1(1+\theta)(t-1)}{2p(1+\theta)} = z_1 - u$. So $H(z)$ is continue in z_1

Third, for z_3 defined by $G_1(e_1^*(z_3)) = G_2(z_3 - u)$, so $H(z)$ is continue in z_3.

CASES AROUND u

$$z_1^* = u \leftrightarrow A = A_0 = \frac{u}{1-t_0}$$
$$z_2^* = u \leftrightarrow A = A_0' = \frac{u}{1 - up_0(1+\theta)(t_1 - t_0)}$$

$A > A_0$ because we assume $u > \frac{1}{p_0(1+\theta)(t-1)}$. Also $H_1(u) > H_2(u)$

Case I: $A < A_0' \leftrightarrow z_2^* < u$ & $A > A_0 \leftrightarrow z_1^* > u$

$$\max[H_1(u), H_2(u)] = H_1(u)$$

Case II: $A < A_0' \leftrightarrow z_2^* < u$ & $A < A_0 \leftrightarrow z_1^* < u$

$$\max[H_1(z_1^*), H_2(u)] = H_1(z_1^*)$$

Because $H_1(z_1^*) > H_1(u) > H_2(u)$

Case III: $A > A_0' \leftrightarrow z_2^* > u$ & $A > A_0 \leftrightarrow z_1^* > u$

$$\max[H_1(u), H_2(z_2^*)]$$

Lead to a quadratic equation with one solution greater than A_0':

$$\begin{aligned}
A_1 = \Big(&t_0 + 4p_0u + t_0^{1/2}(t_0 + 8p_0u + 8p_0^2t_0u^2 + 8p_0^2t_1u^2 + 8p_0\theta u \\
&+ 8p_0^2t_0\theta^2u^2 + 8p_0^2t_1\theta^2u^2 + 16p_0^2t_0\theta u^2 + 16p_0^2t_1\theta u^2)^{1/2} \\
&+ 4p_0^2t_0u^2 - 4p_0^2t_1u^2 + 4p_0\theta u + 4p_0^2t_0\theta^2u^2 - 4p_0^2t1\theta^2u^2 \\
&+ 8p_0^2t_0\theta u^2 - 8p_0^2t_1\theta u^2\Big)\Big/\Big[4p_0(\theta+1)(p_0^2t_0^2\theta^2u^2 + 2p_0^2t_0^2\theta u^2 \\
&+ p_0^2t_0^2u^2 - 2p_0^2t_0t_1\theta^2u^2 - 4p_0^2t_0t_1\theta u^2 - 2p_0^2t_0t_1u^2 \\
&+ p_0^2t_1^2\theta^2u^2 + 2p_0^2t_1^2\theta u^2 + p_0^2t_1^2u^2 + 2p_0t_0\theta u + 2p_0t_0u - 2p_0t_1\theta u - 2p_0t_1u - t_0t_1 + 1)\Big]
\end{aligned}$$

In conclusion $A_0 < A_0' < A_1$ and

$$z^* = \begin{cases} z_1^* & \text{if } A \le A_0 \\ u & \text{if } A \in (A_0, A_1) \end{cases} \tag{11}$$

CASES AROUND $\bar{z}$

$$z_3^* = \bar{z} \leftrightarrow A = A_3 = \frac{\bar{z}}{1-t_1}$$
$$z_2^* = \bar{z} \leftrightarrow A = A_3' = \frac{\bar{z}}{1 + p_0 u(1+\theta)(t_1+t_0) - 2\bar{z}p_0 t_1(1+\theta)}$$

$A_3 < A_4$. And $H_2(\bar{z}) = H_3(\bar{z})$ because of continuity of H in $\bar{z}$

Case I $A < A_3 \leftrightarrow z_3^* < \bar{z}$ & $A < A_3' \leftrightarrow z_2^* < \bar{z}$

$\max[H_2(z_2^*), H_3(\bar{z})] = H_2(z_2^*)$

Case II $A > A_3 \leftrightarrow z_3^* > \bar{z}$ & $A > A_3' \leftrightarrow z_2^* > \bar{z}$

$\max[H_2(\bar{z}), H_3(z_3^*)] = H_3(z_3^*)$

Case III $A > A_3 \leftrightarrow z_3^* > \bar{z}$ & $A < A_3' \leftrightarrow z_2^* < \bar{z}$

$\max[H_2(z_2^*), H_3(z_3^*)]$

Lead to a quadratic equation with root:

$$\begin{aligned}
A_2 = \Big(&2t_1 + 2t_1\theta - 2t_1^2\theta - 2t_1^2 + 2p_0t_0u + 2p_0t_1u + p_0^2t_0^2u^2 \\
&+ p_0^2t_1^2u^2 - 4p_0t_0t_1u + 4p_0t_0\theta u + 4p_0t_1\theta u + 3p_0^2t_0^2\theta u^2 + 3p_0^2t_1^2\theta u^2 + 2p_0t_0\theta^2u + 2p_0t_1\theta^2u \\
&+ 3p_0^2t_0^2\theta^2u^2 + p_0^2t_0^2\theta^3u^2 + 3p_0^2t_1^2\theta^2u^2 + p_0^2t_1^2\theta^3u^2 + p_0^{1/2}u^{1/2}(\theta+1)^{3/2}(-(t_0-t_1)(2t_1 - p_0t_0u \\
&+ p_0t_1u - p_0t_0\theta u + p_0t_1\theta u)(p_0^2t_0^2\theta^2u^2 + 2p_0^2t_0^2\theta u^2 + p_0^2t_0^2u^2 - 2p_0^2t_0t_1\theta^2u^2 - 4p_0^2t_0t_1\theta u^2 \\
&- 2p_0^2t_0t_1u^2 + p_0^2t_1^2\theta^2u^2 + 2p_0^2t_1^2\theta u^2 + p_0^2t_1^2u^2 \\
&- 6p_0t_0t_1\theta u - 6p_0t_0t_1u + 4p_0t_0\theta u + 4p_0t_0u - 2p_0t_1^2\theta u - 2p_0t_1^2u \\
&+ 4p_0t_1\theta u + 4p_0t_1u - 4t_1 + 4))^{1/2} - 2p_0^2t_0t_1u^2 - 4p_0t_0t_1\theta^2u - 6p_0^2t_0t_1\theta u^2 - 8p_0t_0t_1\theta u \\
&- 6p_0^2t_0t_1\theta^2u^2 - 2p_0^2t_0t_1\theta^3u^2\Big)/\Big(4p_0t_1(t_1-1)^2(\theta+1)^2\Big)
\end{aligned}$$

In conclusion, $A_1 < A_3 < A_2 < A_3'$ and:

$$z^* = \begin{cases} z_2^* & \text{if } A \in (A_1, A_2) \\ z_3^* & \text{if } A > A_2 \end{cases} \tag{12}$$

From (10), (11) and (12):

$$z^* = \begin{cases} (1-t_0)A & \text{if } A < A_0 \\ u & \text{if } A \in [A_0, A_1) \\ \frac{A[1+up_0(1+\theta)(t_1+t_0)]}{1+2p_0t_1A(1+\theta)} & \text{if } A \in [A_1, A_2) \\ (1-t_1)A & \text{if } A > A_2 \end{cases} \tag{13}$$

$$e^* = \begin{cases} \frac{1}{2p_0(1+\theta)} & \text{if } A < A_1 \\ z - u & \text{if } A \in [A_1, A_2) \\ \frac{1}{2p_0(1+\theta)} & \text{if } A > A_2 \end{cases} \tag{14}$$

Where:

$$A_0 = \frac{u}{1 - t_0}$$

$$A_1 = \Big(t_0 + 4p_0 u + t_0^{1/2}(t_0 + 8p_0 u + 8p_0^2 t_0 u^2 + 8p_0^2 t_1 u^2 + 8p_0 \theta u$$
$$+ 8p_0^2 t_0 \theta^2 u^2 + 8p_0^2 t_1 \theta^2 u^2 + 16p_0^2 t_0 \theta u^2 + 16p_0^2 t_1 \theta u^2)^{1/2}$$
$$+ 4p_0^2 t_0 u^2 - 4p_0^2 t_1 u^2 + 4p_0 \theta u + 4p_0^2 t_0 \theta^2 u^2 - 4p_0^2 t1 \theta^2 u^2$$
$$+ 8p_0^2 t_0 \theta u^2 - 8p_0^2 t_1 \theta u^2\Big) / \Big[4p_0(\theta + 1)(p_0^2 t_0^2 \theta^2 u^2 + 2p_0^2 t_0^2 \theta u^2$$
$$+ p_0^2 t_0^2 u^2 - 2p_0^2 t_0 t_1 \theta^2 u^2 - 4p_0^2 t_0 t_1 \theta u^2 - 2p_0^2 t_0 t_1 u^2$$
$$+ p_0^2 t_1^2 \theta^2 u^2 + 2p_0^2 t_1^2 \theta u^2 + p_0^2 t_1^2 u^2 + 2p_0 t_0 \theta u + 2p_0 t_0 u - 2p_0 t_1 \theta u - 2p_0 t_1 u - t_0 t_1 + 1)\Big]$$

$$A_2 = \Big(2t_1 + 2t_1\theta - 2t_1^2\theta - 2t_1^2 + 2p_0 t_0 u + 2p_0 t_1 u + p_0^2 t_0^2 u^2$$
$$+ p_0^2 t_1^2 u^2 - 4p_0 t_0 t_1 u + 4p_0 t_0 \theta u + 4p_0 t_1 \theta u + 3p_0^2 t_0^2 \theta u^2 + 3p_0^2 t_1^2 \theta u^2 + 2p_0 t_0 \theta^2 u + 2p_0 t_1 \theta^2 u$$
$$+ 3p_0^2 t_0^2 \theta^2 u^2 + p_0^2 t_0^2 \theta^3 u^2 + 3p_0^2 t_1^2 \theta^2 u^2 + p_0^2 t_1^2 \theta^3 u^2 + p_0^{1/2} u^{1/2} (\theta + 1)^{3/2} (-(t_0 - t_1)(2t_1 - p_0 t_0 u$$
$$+ p_0 t_1 u - p_0 t_0 \theta u + p_0 t_1 \theta u)(p_0^2 t_0^2 \theta^2 u^2 + 2p_0^2 t_0^2 \theta u^2 + p_0^2 t_0^2 u^2 - 2p_0^2 t_0 t_1 \theta^2 u^2 - 4p_0^2 t_0 t_1 \theta u^2$$
$$- 2p_0^2 t_0 t_1 u^2 + p_0^2 t_1^2 \theta^2 u^2 + 2p_0^2 t_1^2 \theta u^2 + p_0^2 t_1^2 u^2$$
$$- 6p_0 t_0 t_1 \theta u - 6p_0 t_0 t_1 u + 4p_0 t_0 \theta u + 4p_0 t_0 u - 2p_0 t_1^2 \theta u - 2p_0 t_1^2 u$$
$$+ 4p_0 t_1 \theta u + 4p_0 t_1 u - 4t_1 + 4))^{1/2} - 2p_0^2 t_0 t_1 u^2 - 4p_0 t_0 t_1 \theta^2 u - 6p_0^2 t_0 t_1 \theta u^2 - 8p_0 t_0 t_1 \theta u$$
$$- 6p_0^2 t_0 t_1 \theta^2 u^2 - 2p_0^2 t_0 t_1 \theta^3 u^2\Big) / \Big(4p_0 t_1 (t_1 - 1)^2 (\theta + 1)^2\Big)$$

8.2 Estimation of the model only with SII data

We could assume the data from the appraisal company has significant measure of error and therefore is not useful. Nevertheless we can do the method of moments based only on data from the SII and use our model to estimate values for the actual size of the properties. We re-estimate the simulated method of moments with the following moments:

$$\hat{m}_1 = \frac{\sum_{i=1}^{N} 1(d(i) \in [60, 90))}{N}$$
$$\hat{m}_2 = \frac{\sum_{i=1}^{N} 1(d(i) \in [90, 135))}{N}$$
$$\hat{m}_3 = \frac{\sum_{i=1}^{N} 1(d(i) \in [135, 140])}{N}$$
$$\hat{m}_4 = \frac{\sum_{i=1}^{N} 1(d(i) \in (140, 240])}{N}$$
$$\hat{m}_5 = \frac{\sum_{i=1}^{N} 1(d(i) > 240))}{N}$$

The next tables and figures (Tables 6 and 7, Figure 11) present the results of the method of moments. The fit is good for the distribution of d and the conclusions from the counterfactual exercise maintain.

Table 6. Parameters

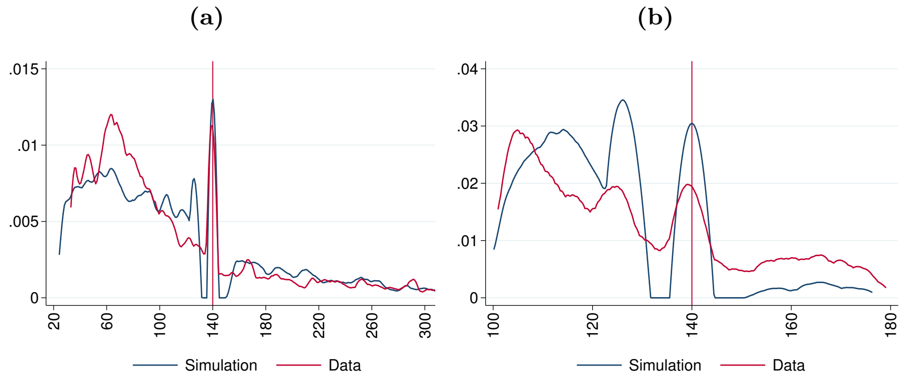
	$\alpha = 0$
μ_A	4.70
	(0.014)
σ_A	0.65
	(0.010)
p_0	3.48%
	(0.09%)

Note: Bootstrap standard deviations in parenthesis.

Table 7. Data and estimated moments

	Data	$\alpha = 0$
Average absolute deviation	–	0.0864
m_1 ($d \in [60, 90)$)	25.2%	19.8%
m_2 ($d \in [90, 138)$)	20.9%	20.6%
m_3 ($d \in [135, 140]$)	7.1%	7.3%
m_4 ($d \in (140, 240]$)	13.2%	16.2%
m_5 ($d > 240$)	10.0%	9.4%

Figure 11. Kernel density of data and simulated model II: d



Note: Epanechnikov kernel; half-width of kernel = 2.

We also developed an estimation of the model where z had measurement of error (not shown in this paper). Nevertheless, this extension impose a trade-off between the fit of the bunching of z and the proportion of households with negative underreporting. In the end, this extensions was not beneficial for our purposes.