Present Bias and Self-Control: Structural Estimation from a Mortgage Market in Mexico

By Emilio Cuilty*

Draft: November 6, 2017

This article studies present bias and self control in a real mortgage market. I look at a novel dataset from a Mexican mortgage institution, where I individually follow sate workers repayment decisions across 15 years. By applying Fang and Wang (2016) methodology, I am able to estimate the long term discount factor δ, the quasi-hyperbolic discount factor β and the degree of naivety $\tilde{\beta}$ the state worker has. I find that the state worker suffers from present bias ($\beta = 0.35$), and is not aware of it ($\tilde{\beta} = 1$). I show that the mortgage debt could be repaid faster if individuals could behave either as exponential discounters, or as sophisticated present biased discounters. These finding suggest that the risk of default is greater than the one estimated under traditional exponential models.

It is not surprising that long term decisions require deeper levels of rationality by individuals. Retirement savings, mortgages, and credit loans are examples where agents wrongly estimating their own future payoffs. From a behavioral economics perspective, this phenomenon has been explained by bounded rationality, reweighting of outcome probabilities, loss aversion, present bias and lack of self-control[1]. Although most of these theories have been tested in an experimental framework, there is a shortage of non-experimental empirical work (Andreoni and Sprenger 2012). Therefore some implications

* Cuilty: University of Wisconsin Madison, Department of Economics, Observatory Drive, Madison, WI 05715 (e-mail: cardenascuil@wisc.edu). Acknowledgments: This research relied heavily on data provided by FOVISSSTE through the Central Bank of Mexico “Banxico”. This inter-institutional cooperation was possible due to the efforts of José Reyes Baéza, Hector Valles and Edgar Cortes. I appreciate the comments made in the SABE/IAREP conference, specially those from David Leiser. Finally this work would never existed without the support and guidance of Alan Sorensen.

[1] A good survey can be found in DellaVigna (2009), Oshry (2006) or Robson and Samuelson (2009) .

of these experimental findings might not be applicable for policy making, mechanism design or expanding the economics core theory. It is no coincidence that there is a lack of non-experimental research, as noted by de Clippel and Rozen (2014), as testing theories that violate some rationality assumptions require large data sets of outstanding quality. Nevertheless, the growing availability of individual data presents a good opportunity for doing so.

By looking into a real credit market, the present paper aims to test one of the most popular theories on present bias and self-control: Quasi-Hyperbolic Discounting (QHD) [2]. An individual that exhibits QHD places additional weight on present payoffs over long run payoffs, i.e. he is present biased. Consequently, naive present biased individuals borrow excessively and often fail to repay debt (Kuchler 2015). If banks or any other credit institutions fail to take the effects of QHD into consideration, they might underestimate the risk of default. In developing countries where the credit market is relatively small this could be extremely problematic.

To test QHD I analyze the debt repayment behavior of state workers in one of Mexico's public mortgage institutions. Despite the fact that Mexico's mortgage market contains private institutions (such as banks and a small number of housing providers), the public institutions account for 70 percent of the market share (SHF 2014). Among the public sector there exist only two institutions, one for state workers (20 percent of the market share) and one for non-state formal workers (50 percent of the market share). One notable difference between these public agencies is their repayment schemes; in particular their definition of mortgage debt. The state workers have to pay a fixed amount of minimum wages whereas the non-state workers have to pay a fixed amount of Mexican pesos. Facing a mortgage debt measured in minimum wages implies that the real debt increases every time there is a minimum wage increment[3]. Thus a state worker has additional incentives to repay sooner rather than later, but present bias might preclude this behavior.

In this novel data I observe a random sample of state workers over a long period of time, in which it is possible to see how they decide to repay their mortgage. Every worker faces a mortgage contract that fortnightly discounts 30 percent of her base salary as payment, and

[2]While others theories might be interesting as well, QHD has been studied deeper in theoretical and experimental models since its first appearance in 1997 in Laibson's paper.

[3]In Mexico the minimum wage is raised every year, during my sample period said raise was above the inflation rate

she has the option of making extra contributions in order to repay her debt faster. This option of making additional payments allows me to build a dynamic discrete choice model, in which the worker decides every period whether to make an additional contribution or not. It is worth mentioning that I am assuming that the amount of an extra payment is optimal, in this way I only consider the decision of making an extra payment.

Applying the methodology designed by Fang and Wang (2015) I am able to identify present bias, and in particular, if the agents are aware of having this bias. This is achieved by structurally estimating three discount factors (long run, quasi-hyperbolic, and degree of naiviety). Identifying all discount factors relies on using exclusion variables. These variables affect the transition probabilities but not the difference of instantaneous payoffs.[4]. Distinguishing if individuals are sophisticated or naive is important when making public policy recommendations. For example, a sophisticated agent with present bias who has access to an illiquid asset will be able to smooth consumption (Laibson 1997), whereas the naive agent might struggle to do so. Although the present bias hypothesis can be studied from a reduced form approach (see Kuchler (2015)), the structural estimation allows me to conduct simulations of several public policies after recovering the relevant parameters.

For the purpose of this paper Mexico is quite interesting, since it is not yet a fully developed nation and the credit market is relatively small. To my knowledge this is the first paper that uses mortgages to recover the structural parameters of a dynamic discrete choice model. Therefore, this research contributes to the literature of behavioral economics, especially to that which uses structural empirical methods, and it also provides public policy recommendations for Mexico's case. The rest of the paper is organized as follows: Section 1 relates my research to the current literature; the mortgage contract and data are described in Section 2; Section 3 introduces the model; Section 4 explains the econometric strategy as well as the identification argument; the main results and simulations are presented in Section 5; and a closing discussion is provided in Section 6.

I. Relationship with the Literature

The first model with present bias captured by QHD's preferences was introduced in Laibson (1997). Some of the key features of the QHD consumption-savings model are:

[4]For example, gender affects the probability of getting a raise, but it doesn't affect the difference of utilities between making a contribution or not.

aggregate consumption tracks income, present biased individuals save less than dynamically consistent individuals, and in some situations, defaulting might be optimal. An important prediction of this model is that if consumers have access to an illiquid asset (i.e., a commitment device) present biased individuals are able to smooth consumption[5]. Without modeling QHD explicitly, this theoretical feature has been well documented in several experiments. For example, by offering an illiquid saving program, Thaler and Benartzi (2004) observed an increase of 10 percent in individual savings. Similarly but varying the intensity of the commitment, Dupas and Robinson (2013) offered savings accounts to poor people in Kenya. By making available safe accounts, the authors observed an increase in savings. Moreover health emergencies were covered thanks to iliquid characteristics of some of the accounts. When social pressure was added to the mix, the saved amount increased substantially. Further discussion and classification of commitment devices can be found in Bryan et al. (2010). The authors remark that psychological punishments, enormous penalties fees, and blocking access to accounts might have different effects on how individuals allocate their resources. Importantly the mortgage contract analyzed in this research, discounts 30 percent of the agents base salary, imposing some sort of commitment.

The first empirical research that modeled QHD explicitly comes from representative agent models with aggregate variables. For instance Angelatos et al. (2001) showed that a QHD model approximates aggregate data patterns better than an exponential discount model. By calibrating the discount parameters, the authors explain why households have very little liquid assets and at the same time maintain a substantial amount of debt in their credit cards. As shown in the work of Laibson et al. (2017), the former findings are robust to modern econometric techniques such as the simulated method of moments. Moreover the authors show that a QHD model can replicate additional stylized facts from the U.S. data, such as retirement wealth accumulation. Combining both aggregate and individual data, Dellavigna and Paserman (2005) tested theoretical implications of present bias in a job search model. Most of the unemployment data facts are only consistent with the predictions of their QHD model, rejecting the exponential hypothesis. As it is common in structural work, aggregate data questions the identification of the relevant parameters. It will be shown later that having individual data provides a cleaner way of estimating the

[5] While Laibson's model only considers illiquid assets as commitment devices, it is possible to find several other devices with the same effect.

discount parameters.

As is common in behavioral economics several papers conducted experiments to explore present bias and in particular QHD. In a university experiment Ariely and Wertenbroch (2002) found that present bias explains why individuals fail to finish their homework on time. It is not just about procrastination, present bias matters in credit markets. Meier and Sprenger (2010) found that present biased individuals are more likely to borrow more than dynamically consistent individuals. This finding comes by combining information from a choice experiment (to identify time preferences) and real credit card usage data. Jones and Mahajan (2015) carried out a field experiment in which they were able to identify both the long run discount and the present bias discount in a QHD structural model. Their experiment varied exogenously the time and amount of money an individual would receive in a real tax refund. Such variation can hardly appear outside the experimental framework, and therefore identification of structural parameters requires additional assumptions.

Among the papers that use non-experimental individual data, the reduced form approach seems to be more popular. Using a gym attendance data set, DellaVigna and Malmendier (2006) tested some predictions of a model with standard time preferences and a model with present bias. They found that standard intertemporal preferences do not capture the behavior of the data, while QHD models do. The research of Kuchler (2015) analyzed credit card data at an individual level. She had access to a software designed to help individuals repay their credit card debt. By observing individual transactions as well as the individual's desired level of debt, she was able to distinguish between individuals that were not present biased from those who were. Among those who had present bias, she distinguished between those individuals that believed they discounted in an exponential way from those that were aware of being present biased (i.e., naive and sophisticated agents). She rejected exponential discount and found that naive present biased individuals are very likely to default. One drawback of the reduced form approach is the lack of counterfactuals, and thus the inability of simulating public policies.

The structural models were avoided due to the complexity of identifying the relevant discount parameters. Rust (1994) showed that discount factors in a standard dynamic discrete choice model are generically not identified. By introducing exclusion restrictions Magnac and Thesmar (2002) were able to identify the exponential discount factor. In the same way as Hotz and Miller (1993) , the paper of Fang and Wang (2015) solves the identification issue by using exclusion restrictions similar to those by Magnac and Thesmar (2002). The present paper closely follows Fang and Wang (2015) and, like them, I am able to identify the parameter that characterizes the long run discount, the QHD, and the agent's naivety degree.

Several structural papers that rely on parametric assumptions or some exogenous variations are also worth mentioning. Fang and Silverman (2009) implemented a finite horizon dynamic discrete choice model and analyzed a welfare program. Their identification results rely on using a large panel data set with a clear final period. They showed that there are no observationally equivalent set structures with exponential discounting that yield QHD predictions[6]. However they could not identify the naivety degree. Paserman (2008) analyzed a job search model in which identification is achieved by the large variation in unemployment spells and accepted wages. Although he rejected exponential discounting, his results depend on the model's specific structure and on the functional form of the wage distribution.

Finally some axiomatic derivations to test present bias have been made in the paper of Echenique et al. (2017). Using experimental data their test showed that half of their sample was consistent with exponential discounting and only a few more were consistent with QHD. The test consists in a non-parametric revealed preference approach, nevertheless to implement it several choice decisions need to be observed. This escapes the nature of the data used in this paper.

II. Background and Data description

The data come from a de-centralized Mexican public mortgage institution, *Fondo de la Vivienda del Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado* (State Worker's Social Security and Services Housing Fund), hereafter FOVISSSTE. FOVISSSTE was created in 1972 in order to meet the housing credit demand for most of Mexico's state workers (approximately 2.3 millions workers). The state workers affiliated to FOVISSSTE are either federal, state and municipal governments, as well as public universities and local agencies. FOVISSSTE offers programs with different credit schemes, I will use a sample from the most important one, the "Traditional Credit Scheme"[7].

[6] They used National Longitudinal Surveys from 1979.

[7] The Traditional Credit Scheme comprehends somewhere around 62 percent of their total mortgages.

The Traditional Credit Scheme works like this: Every year FOVISSSTE allows workers to register in a lottery in which some fixed number of mortgages are randomly assigned (in 2013, 45,000 mortgages were given). If the worker is selected, she has to accept a mortgage contract that offers no more than U.S. $58,852[8]. The mortgage's value depends on the worker's base salary, and it should be fully repaid in 30 years. After signing the contract the worker chooses a house from a menu of external companies affiliated to FOVISSSTE. Once she selects a house, FOVISSSTE transfers the money to the external company, and when she receives her house she starts repaying the mortgage. Each fortnight FOVISSSTE charges a fixed interest rate. The assigned interest rate depends on her base salary (it could go from 4 percent to 6 percent yearly), and it is applied to the current outstanding debt.

I consider a random sample of 410 workers that start repaying their mortgage from the first week of January 2000 to the last week of June 2015. Every two weeks I observe the transactions that both the individuals (payments) and FOVISSSTE (charges) have made. Therefore I have a panel of $T = 348$ periods with potentially 142,000 observations[9]. If the worker is employed I observe her base salary, otherwise I assume the worker has no income. It's worth mentioning that I am not able to find other sources of income which potentially can bias my results, hereafter I will use wage, salary and income indistinctly.

Time invariant statistics of my sample are presented in Table 1. When the individuals enter the panel their age ranges from 21 to 62 years old, where the average worker is 38 years old. Among the workers, 43 percent live in Mexico City and 61 percent of my sample are women. The average interest rate is 5.42 percent.

TABLE 1—TIME INVARIANT STATISTICS

Variable	Mean	Standard Deviation	Min	Max
Gender (Female==1)	0.61	0.4872	-	-
Age at 2000	38.05	7.8345	21	62
Interest Rate	5.42	1.4566	4	6
Mexico City	0.43	0.4959	-	-

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000.

The repayment scheme works like this: FOVISSSTE fortnightly deducts 30 percent of the

[8]As of the writing, the exchange rate is approximately $16 pesos per dollar.
[9]My data set consists on 110,356 observations, since any agent who repays the the full mortgage exits the panel.

worker's base salary. If the worker is unemployed, it is possible to make contributions to the mortgage, however in the data this is a rare scenario[10]. Every two months the agency where she works makes an extra 5 percent contribution of her base salary. Additional payments can be done by the worker every fortnight, these payments can be of any magnitude. Recall that the mortgage is measured in minimum wages, whenever there is an increment in the minimum wage the oubstanding real debt increases. In my sample period every January the minimum wage increased on average 7 percent[11].

A. Mortgage Debt

Before analyzing some insightful characteristics of the mortgage debt it would be useful to introduce its law of motion. Let $S \subset T$ be the set of fortnights when the additional 5 percent contribution occurs, and let $W \subset T$ be the set of fortnights in which the minimum wage increment takes place. Then in every fortnight $t \in T$, whenever $m_{t-1} > 0$ the law of motion of the mortgage debt is captured by:

$$m_t = [1 + \Delta_t \mathbf{1}\{t \in W\}]\,[(1+r)m_{t-1}] - .3y_t - a_t - .05y_t \mathbf{1}\{t \in S\} \tag{1}$$

where m_t is the value of the mortgage at time t, y_t is the base salary of the worker at time t, r is the interest rate , a_t is the additional contribution and Δ_t is the minimum wage's increment.

Table 2—Mortgage Debt Statistics

Year	Mean	Median	S.D.	State Workers
1	23,069	21,952	6,605	410
7.5	16,581	15,142	6,063	370
15	9,989	8,548	7,014	316

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. Mortgage debt is measured in real U.S. dollars, in which 1 Mexican peso equals 16 dollars. The year 2015 is used as base year.

Table 2 presents mortgage debt statistics at different periods of my sample. On average both the mean and median debt decreased over time, which is expected given the nature

[10] Only .3 percent of my sample made a contribution when unemployed.
[11] From January 2000 to January 2015 the minimum wage has increased 82 percent.

of how the debt is repaid. Nevertheless it is possible to observe that some workers finished paying their mortgage within 15 years. Also note that the variance is relatively higher as time passes. In other words my sample is heterogeneous regarding debt repayment. One possible explanation is that some workers experienced meaningful increments in their wage over time, it also could be that some workers decided to pay more than just 30 percent of their income.

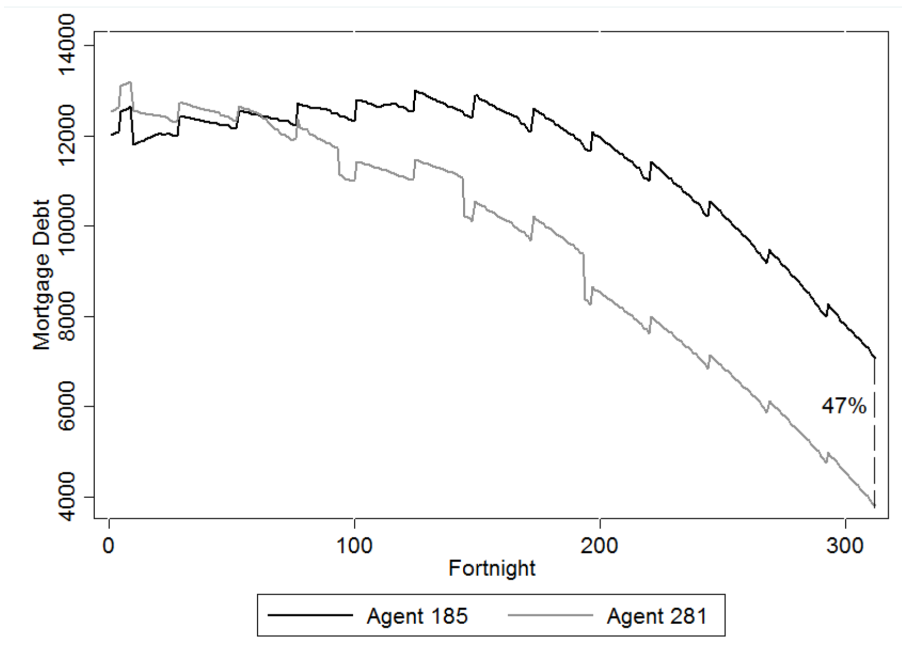

Figure 1. Individual Mortgage Comparison

Note: Individual mortgage debt across time. All variables are measured in U.S. dollars, where $16 Mexican pesos equals $1 dollar.

To illustrate how different repayment behavior changes the mortgage debt over time I show the comparison between two state workers in figure 2.1. Both state workers started with very similar mortgage debt, yet there is a 47 percent difference after 15 years. In this figure it is possible to observe both the minimum wage increments (cyclical upward jumps) and some additional contributions (downward jumps), the workers' income is captured by the slope of the series. My hypothesis is that because the debt increases every year and, since it is possible to lose the job in such a long period span, a worker has incentives to repay as soon as possible. Nevertheless present bias can preclude this behavior. Thus if both workers in the series have similar income patterns, the 47 percent difference might be explained by present bias.

B. Income

Some standard economic theories predict that income grows when we consider a long period span. Table 3 depicts the former fact in the first 7 years, both the median and the average worker experienced real income growth. On the other hand the last 8 years of my sample remain statistically the same. Therefore the hypothesis that individuals should make additional contributions as soon as possible might not be true for the first years, since the worker might be expecting wage increments. Still if income remains constant (as for the last 8 years) , or if it is possible to experience a job loss, an exponential discounter should make early additional contributions.

Table 3—Annual Base Salary Statistics

Year	Mean	Median	S.D.	State Workers
1	5,056	4,635	2,880	410
7.5	7,136	5,336	6,039	370
15	7,043	5,536	5,847	316

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. The annual base salary is measured in U.S. dollars, where 1 Mexican peso equals 16 dollars. The year 2015 is used as a base year.

Figure 2.2 presents a comparison of the same state workers as in last figure, but now comparing income. First note that real wage is considerably stable over time. We can observe some temporary income changes due to penalization at work (i.e., late arrivals or missing days) or small bonus due to extra hours worked that fortnight. Note that both workers have a very similar income pattern with a correlation of .89. In other words, the 47% difference observed in the mortgage debt is not coming from income shocks but from additional payments.

Table 4 shows how income varies across different groups. On average male state workers earn more than women, previous studies have shown that Mexico suffers from gender discrimination in terms of wages (Arceo-Gomez and Campos-Vazquez 2014). Predictably workers older than 40 earn more than young workers. Living in Mexico City, on the other hand, implies having less income than in the rest of the country. These variations across

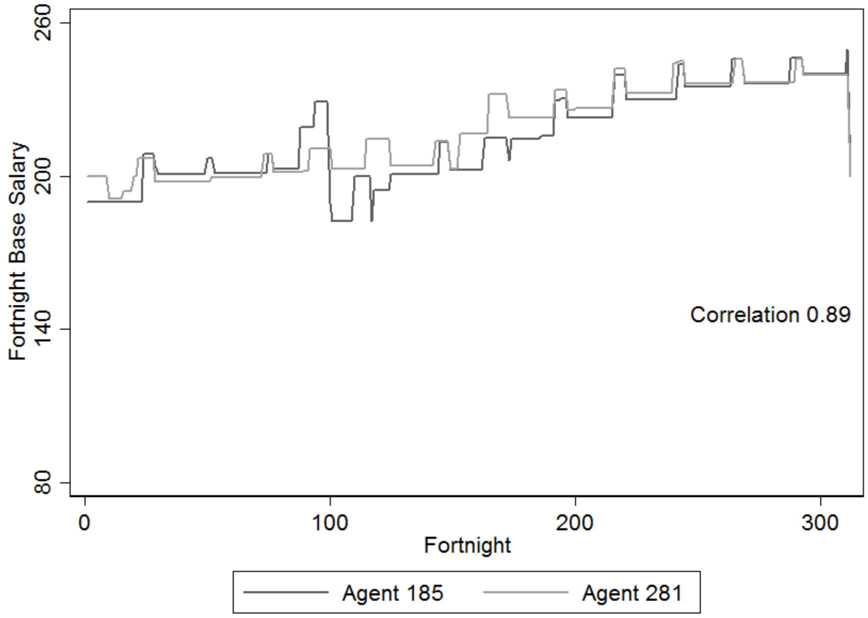

FIGURE 2. INDIVIDUAL BASE SALARY COMPARISON

Note: Income across time. U.S. dollars, where 1 Mexican peso equals 16 dollars. Income is considered to be the fortnight base salary.

groups will be relevant at the time of estimating transition probabilities, which will be detailed later in the econometric specification.

C. Additional Payments

As discussed earlier additional payments can have a large impact on how fast the mortgage debt is repayed. Table 5 presents the average additional payment across different years in my sample. Since several workers never made an additional contribution, I also provide the mean that excludes those workers. Overall the largest payments were made during the first years. These early payments might lead some workers to exit my sample. Note that when excluding the zero additional payment the average contribution increases over time. The opposite effect occurs with the frequency of workers. In other words, as time goes we have less workers who decided to contribute, and among them, the payment is higher. As hypothesized earlier, present bias might preclude workers from making early contributions. To have the same effect on the mortgage debt, due to minimum wage increments, latter

Table 4—Average Income Across Groups

	Income mean	Standard Deviation
Female	6,140	6,893
Male	6,893	9,232
Mexico City	6,051	5,958
Rest of the Country	6,698	7,538
Age<40	5,636	5,617
Age>40	7,423	8,188

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. All variables are measured in U.S. dollars, where 1 Mexican peso equals 16 dollars. Income is considered to be the base salary a worker earned in a year. The year 2015 is used as a base year.

payments need to be greater than early payments. Workers that never supply additional contributions might not be able to finish paying their mortgage in 30 years, and so, they will enter into default.

Table 5—Annual Additional Payments Statistics

Year	Mean	S.D.	Mean Excluding Zero	S.D. Excluding Zero	State Workers	Workers with Extra Payment
1	185.12	504.90	703.09	864.43	410	95
7.5	135.60	604.68	1,570.83	1,868.91	370	34
15	158.28	956.74	2,278.41	3,307.59	316	19

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. Annual payment is measured in U.S. dollars, where 1 Mexican peso equals 16 dollars. The year 2015 is used as base year.

III. The Model

In this section I present a dynamic discrete choice model that substantiates the optimization problem of a state worker with a FOVISSSTE mortgage. Both the model and the identification argument will use the results developed in the paper of Fang and Wang (2015), extensive derivations and proofs can be found in their paper. Before appealing to the discrete choice framework it will be useful to present a continuous choice model. This underlying model motivates how state variables will be defined in the discrete choice approach.

A. The Underlying Model: Continuous Choice

Assume that time horizon is infinite, indexed by $t = 1, 2, 3...$, where every period represents a fortnight. The state worker's preferences are defined over a sequence of consumption and random unobserved preference shocks $U_t : \{c_t,\, \varepsilon_t\}_{t=1}^{\infty} \to \mathbb{R}$. Such time preferences have the same (β, δ)-preferences as Laibson (1997).

$$U_t(u_t, u_{t+1}, ...) = u^*(c_t,\, \varepsilon_t) + \beta \sum_{k=t+1}^{\infty} \delta^{k-t} u^*(c_k,\, \varepsilon_t) \tag{2}$$

Where $u^*(\cdot)$ is the instantaneous utility, $\beta \in (0, 1]$ captures present bias, referred also as the present bias discount or QHD discount, and $\delta \in (0, 1]$ is the long run discount. Note that if $\beta = 1$ the model becomes the classic exponential discount model.

The state worker faces a mortgage contract that deducts automatically 30 percent of her income. Every period the worker chooses how much to consume and implicitly how much additional payment to make, there are no extra savings in the model. Therefore the budget constraint that the state worker faces every fortnight t is defined by:

$$c_t + a_t \mathbf{1}\{m_t > 0\} \leq .7y_t + .3y_t \mathbf{1}\{m_t = 0\} \tag{3}$$

The budget constraint depends on the value of the mortgage debt m_t , and so the worker needs to consider its law of motion defined in equation (1). Note that the law of motion relies on the increment of the minimum wage Δ_t, hence the worker is uncertain about the value of her mortgage debt in some periods. In particular let the stochastic process of the minimum wage increment be described as

$$\Delta_t = \Delta_{t-24} + \Omega_t \tag{4}$$

Where Ω_t is an i.i.d. random shock.

The worker is uncertain about her future income as well. Inspired by the former section, the random component will be related to her invariant features. She knows that the stochastic income process is defined as follows:

$$y_t = y_{t-1} + \gamma_t \tag{5}$$

where γ_t is a random variable that depends on the worker's characteristics, such as gender, age and where the worker lives.

Taking into account equations (1), (2), (3), (4) and (5) the problem of the state worker can be understood as

$$\underset{\{a_t\}_{t=1}^{\infty}}{Max}\, u^*(.7y_t - a_t - \varepsilon_t) + \beta \sum_{k=t+1}^{\infty} \delta^{k-t} E\left[u^*(.7y_k - a_k - \varepsilon_t) | m_{k-1},\, y_{t-1}\right]$$

$$s.t.$$

$$m_t = [1 + \Delta_t \mathbf{1}\{t \in W\}]\,[(1+r)m_{t-1} - .05y_{t-1}1\{t \in S\}] - .3y_t - a_t$$

$$a_t \geq 0$$

The solution to this problem will yield an optimal sequence of consumption and additional payments $\{c_t^*, a_t^*\}_{t=0}^{\infty}$. At the end how much utility the worker gets when making an additional contribution will depend on the values of state variables such as: income y_t, the random preference shock ε_t and the outstanding debt after charges, $M_t \equiv [1 + \Delta_t \mathbf{1}\{t \in W\}]\,[(1+r)m_{t-1} - .05y_{t-1}1\{t \in S\}]$. These state variables, and the way they relate to invariant characteristics, motivate the construction of the discrete choice model.

B. The Discrete Choice Model

The final goal of the paper is to obtain estimates of the discount parameters. To do so I will rely on the discrete choice framework. Instead of considering the value of a_t, assume that the agent decides whether to make an optimal additional payment or no additional payment at all, $i \in \mathcal{I} = \{0, \mathbf{1}\{a_t = a_t^*\}\}$. In this way the problem is transformed into a dynamic discrete choice model. Now it is possible to define the preferences over the discrete choice $i \in \mathcal{I}$, state variables $x_t \in \mathcal{X}$, and a choice specific preference vector shock $\boldsymbol{\varepsilon_t} = (\varepsilon_{0t}.\varepsilon_{1t})$.

$$U_t(u_t, u_{t+1}, ...) = u_i^*(x_t, \boldsymbol{\varepsilon}_t) + \beta \sum_{k=t+1}^{\infty} \delta^{k-t} u_i^*(x_k, \boldsymbol{\varepsilon}_k) \tag{6}$$

Motivated by the underlying model the set of state variables $\mathcal{X}$ contains: the outstanding debt, her income level and some personal characteristics such as age, gender and where does the state worker lives. The choice specific vector $\boldsymbol{\varepsilon_t}$ changes how the state worker feels if she makes an additional contribution. For example if on a given fortnight t the worker has access to a better quality consumption bundle[12], her happiness level when making an additional contribution will be less than when she has access to her regular consumption bundle. For econometric purposes it is useful to make the following assumption:

Assumption 1 ***(Additive separability)***: For each $i \in \{0, 1\}$, the instantaneous utilities are given by,

$$u_i^*(x_t, \boldsymbol{\varepsilon}_t) = u_i(x_t) + \varepsilon_{it} \tag{7}$$

Where $u_i(x_t)$ is the utility's deterministic component from choosing i at x_t, and $(\varepsilon_{0t}, \varepsilon_{1t})$ has a joint distribution G_t which is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^2$.

It is well known that the (β, δ)-preferences may generate dynamic inconsistency. A common approach to account for it is to represent the problem as an interpersonal game. The set of players of the interpersonal game consists in the period-t selves of the same state worker. In accordance with her current utility $U_t(u_t, u_{t+1}...)$, each period-t self decides whether to make or not an optimal additional payment, while her future selves command her subsequent decisions. In this context a Markovian strategy profile for all selves is $\boldsymbol{\sigma} = \{\sigma_t\}_{t=1}^{\infty}$, where $\sigma_t : \mathcal{X} \times \mathbb{R}^2 \to \mathcal{I}$ for all t, i.e, this is a collection of when to make additional payments in all possible states and random shocks. Let $\boldsymbol{\sigma_t^+} \equiv \{\sigma_k\}_{k=t}^{\infty}$ be the continuation strategy profile from period t on, and define the expected continuation utility under her long discount $V_t(x_t, \boldsymbol{\varepsilon}_t; \boldsymbol{\sigma_t^+})$ by

$$V_t(x_t, \boldsymbol{\varepsilon}_t; \boldsymbol{\sigma_t^+}) = u_{\sigma_t(x_t, \varepsilon_t)}(x_t) + \boldsymbol{\varepsilon}_{\sigma_t(x_t, \varepsilon_t)t} + \delta E[V_t(x_{t+1}, \boldsymbol{\varepsilon}_{t+1}; \boldsymbol{\sigma_{t+1}^+}) | x_t, \sigma_t(x_t, \varepsilon_t)] \tag{8}$$

where $\sigma_t(x_t, \varepsilon_t) \in \mathcal{I}$ is the choice specified by σ_t and the expectation is taken over the future states x_{t+1} and future random shocks $\boldsymbol{\varepsilon_{t+1}}$.

As discussed before it is important to consider the possibility that the state worker ignores her own present bias. Like in Stroz (1955), Phelps and Pollak (1968), and O'Donoghue and Rabin (1999) I define a partially naive state worker if her period-t self with present bias β

[12]A better consumption bundle could be sales on given products, access to rare products or events that increase the enjoyment of consumption.

believes her future selves have present bias $\tilde{\beta} \in [\beta, 1]$. In particular if $\tilde{\beta} = \beta$, the period-t self is sophisticated, and if $\tilde{\beta} = 1$ period-t self is completely naive.

The equilibrium will be defined for the partially naive state worker. To that end it is necessary to introduce how the partially naive state workers perceive her future selves will behave.

DEFINITION 1: *(O'Donoghue and Rabin (1999)) A perception continuation strategy profile for a partially naive agent is a strategy profile* $\tilde{\boldsymbol{\sigma}} \equiv \{\tilde{\sigma}_t\}_{t=1}^{\infty}$ *such that for all* $t = 1, 2, ...,$ *all* $x_t \in \mathcal{X}$*, and all* $\boldsymbol{\varepsilon}_t \in \mathbb{R}^2$*,*

$$\tilde{\sigma}_t(x_t, \boldsymbol{\varepsilon}_t) = \underset{i \in \{0, \mathbf{1}\{a_t = a_t^*\}\}}{arg\,max} \left\{ u_i(x_t) + \varepsilon_{it} + \tilde{\beta}\delta E[V_{t+1}(x_{t+1}, \boldsymbol{\varepsilon}_{t+1}; \boldsymbol{\sigma}^{+}_{\boldsymbol{t+1}}) | x_t, i] \right\}$$

Note that definition 1 is not observed in the data but influences the true decision of the state worker. As in Fang and Wang (2015), the equilibrium for a partially naive agent can be defined as a perception-perfect strategy profile.

DEFINITION 2: *A perception-perfect strategy profile for a partially naive agent is a strategy profile* $\boldsymbol{\sigma}^* \equiv \{\sigma_t^*\}_{t=1}^{\infty}$ *such that for all* $t = 1, 2, ...,$ *all* $x_t \in \mathcal{X}$*, and all* $\boldsymbol{\varepsilon}_t \in \mathbb{R}^2$*,*

$$\sigma_t^*(x_t, \boldsymbol{\varepsilon}_t) = \underset{i \in \{0, \mathbf{1}\{a_t = a_t^*\}\}}{arg\,max} \left\{ u_i(x_t) + \varepsilon_{it} + \beta\delta E[V_{t+1}(x_{t+1}, \boldsymbol{\varepsilon}_{t+1}; \tilde{\boldsymbol{\sigma}}^{+}_{t+1}) | x_t, i] \right\}$$

Definition 1 and definition 2 fully characterize the equilibrium of the interpersonal game. Note that when $\tilde{\beta} = \beta$ (i.e., the agent is sophisticated), $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^*$. To apply the identification argument developed in the paper of Fang and Wang (2015) the following assumptions on the data generating process have to be made.

Assumption 2. ***(Stationarity)***: The observed choices are generated under the stationary perception perfect strategy profile of the infinite horizon game played among different selves of the state workers.

Assumption 3. ***(Conditional Independence)***: The transition probabilities satisfy the following:

$$\pi(x_{t+1}, \boldsymbol{\varepsilon}_{t+1} | x_t, \boldsymbol{\varepsilon}_{\boldsymbol{t}}, i_t) = q(\boldsymbol{\varepsilon}_{t+1} | x_{t+1}) \pi(x_{t+1} | x_t, i_t)$$

$$q(\boldsymbol{\varepsilon}_{t+1}|x_{t+1}) = q(\boldsymbol{\varepsilon})$$

Assumption 4. ***(Extreme Value Distribution)***: $\boldsymbol{\varepsilon_t}$ is i.i.d. extreme value distributed.

Imposing stationarity could be unrealistic. A different approach is to estimate the transition probabilities for every period. Note that consistent estimates of the transition probabilities for every fortnight require a large sample of state workers. Since this is not the data's case, I will exploit the long periodicity to estimate the stationary transition probabilities. Obviously it is an assumption that I cannot verify, however since I am considering a long period span, it might be reasonable.

Both assumption 3 and assumption 4 are commonly used in the dynamic discrete choice models. Fang and Wang (2015) remark that it is possible to obtain similar identification results assuming another distribution of the unobservables, however without making any assumption on the distribution the model is not identified. In that sense, making an assumption of G_t imposes a weaker version of semi parametric identification.

At this point it is possible to characterize the decision of the state worker using value functions. Let $W_i(x)$ be the deterministic current choice-specific value function, defined as:

$$W_i(x) = u_i(x) + \beta\delta \sum_{x' \in \mathcal{X}} V(x')\pi(x'|x, i) \tag{9}$$

Where $\pi(x'|x, i)$ is the transition probability from state x to x' when action i is taken, and $V(x)$ is the perceived long run value function, defined as expected value over the stationary value function defined in (8) under the continuation strategy $\tilde{\boldsymbol{\sigma}}$. In some sense equation (9) dictates the behavior of the period-t self.

In the same way let $Z_i(x)$ capture the deterministic perceived choice-specific value function of the future self by the current self as:

$$Z_i(x) = u_i(x) + \tilde{\beta}\delta \sum_{x' \in \mathcal{X}} V(x')\pi(x'|x, i) \tag{10}$$

equation (10) regulates how the current state worker t self believes her future self will behave.

Hence by definition 2 the probability of observing in the data that an additional payment has been made at state x is

$$P_1(x) = Pr[\sigma_t^*(x, \boldsymbol{\varepsilon}) = 1] = Pr[W_1(x) + \varepsilon_1 \geq W_0 + \varepsilon_0] \tag{11}$$

and by assumption 4 it becomes

$$P_1(x) = \frac{exp[W_1(x)]}{exp[W_0(x) + W_1(x)]} \tag{12}$$

Analogously by definition 1 the probability of making an additional contribution by the next period state worker as perceived by the current period state worker at state x is

$$\tilde{P}_1(x) = Pr[\tilde{\sigma}_t(x, \boldsymbol{\varepsilon}) = 1] = Pr[Z_1(x) + \varepsilon_1 \geq Z_0 + \varepsilon_0] \tag{13}$$

and so with assumption 4 it becomes

$$\tilde{P}_1(x) = \frac{exp[Z_1(x)]}{exp[Z_0(x) + Z_1(x)]} \tag{14}$$

Recall that in the data it is possible to observe $P_i(x)$, but $\tilde{P}_i(x)$ is not observable. Finally to fully characterize the model, an expression of $V(\cdot)$ is missing. For this propose let $V_i(x)$ be the perceived choice-specific long run value function as follows:

$$V_i(x) = u_i(x) + \delta \sum_{x' \in \mathcal{X}} V(x')\pi(x'|x, i) \tag{15}$$

Thus using equations (10), (15) and the assumptions on $V(x) = E_\varepsilon[V_{\tilde{\sigma}(x,\varepsilon)}(x) + \varepsilon_{\tilde{\sigma}(x,\varepsilon)}]$, it is possible to get

$$V(x) = ln\left\{exp(Z_0(x) + Z_1(x))\right\} + (1 - \tilde{\beta})\delta \sum_{j=0}^{1} \left[\tilde{P}_j(x) \sum_{x' \in \mathcal{X}} V(x')\pi(x'|x, j)\right] \tag{16}$$

In summary, the dynamic discrete choice model consists on $\{V_i(x), W_i(x), Z_i(x) : x \in \mathcal{X}\}$ as defined by equations (15), (9) and 10), where $W_i(x)$ is the value function that dictates the state worker's behavior in period t; $Z_i(x)$ is what she perceives from her future self choices; and $V_i(x)$ is an auxiliary value function that evaluates payoffs from the choices that the current self perceives will be made by her future selves.

IV. Econometric specification and Identification

In this section I present the econometric specification and some intuition that leads to the identification of the discount parameters $\left\langle \delta, \beta, \tilde{\beta} \right\rangle$. A formal proposition and proof can be found in Fang and Wang (2015).

A. Identification

Denote the structure of the model by θ, identified by the following parameters (Fang and Wang 2015):

$$\theta = \{\left\langle \delta, \beta, \tilde{\beta} \right\rangle,\ G,\ \langle\{u_i(x),\ Z_i(x'),\ V_i(x') : i \in \mathcal{I},\ x \in \mathcal{X},\ x' \in \mathcal{X}\}\rangle\}$$

Where $G(\cdot)$ is a type-I extreme value distribution by assumption 4, and $Z_i(x')$ and $V_i(x')$ satisfy equations (10) and (15), respectively.

Two structures $\theta, \theta' \in \Theta$ are observationally equivalent if the predicted probabilities of making and additional payment by the model are equal, i.e. $\hat{P}_i(x;\theta) = \hat{P}_i(x;\theta')\ \forall i \in \mathcal{I}$ and $x \in \mathcal{X}$. The model is identified if, and only if, for any $\theta, \theta' \in \Theta$, $\theta = \theta'$ if they are observationally equivalent. In other words, there exists no other specification for the utility function, the discount parameters, and the value functions such that they generate the same probabilities. Most of the previous structural present bias modeling requires to assume a parametric form of the utility function, which allows the Rubinstein critique to kick in[13]. The flexibility in the Fang and Wang methodology is robust enough the withstand the Rubinstein critique, at least in observables.

The identification argument of Fang and Wang (2015) consists in two steps. First they show that $\langle\{u_i(x),\ Z_i(x'),\ V_i(x') : i \in \mathcal{I},\ x \in \mathcal{X},\ x' \in \mathcal{X}\}\rangle$ is identified given values of $\left\langle \delta, \beta, \tilde{\beta} \right\rangle$, then they provide conditions to identify the discount parameters. Since the identification argument suggest how to estimate the parameters, I present the basic idea of how each step works.

[13]Rubinstein (2003) shows that choices that are consistent with QHD under a parametric utility function might also be constistent with exponential discounting under a different parametric utility function.

First Step

The first step consists on fixing values of $\left\langle \delta, \beta, \tilde{\beta} \right\rangle$ to identify

$$\langle \{u_i(x),\ Z_i(x'),\ V_i(x') : i \in \mathcal{I},\ x \in \mathcal{X},\ x' \in \mathcal{X}\} \rangle$$

Such identification will be accomplished by solving a system of equations for all states $x \in \mathcal{X}$ and both decisions $i \in \mathcal{I}$. This system of equations will be generated by two blocks.

For the first block let Q be a mapping from $\boldsymbol{W}(x) = (W_0(x), W_1(x))$ to $\boldsymbol{P}(x) = (P_0(x), P_1(x))$. By normalizing one $W_i(x)$ Hotz and Miller (1993) showed that it is possible to invert Q. Using assumption 4, the inverse of this normalized map is $D(x) = W_1(x) - W_0(x) = ln\frac{P_1(x)}{P_0(x)}$. When combining equations (9) and (10) with $D(x)$, it is possible to find a relationship that captures the perceived difference of making an additional payment by future selves:

$$Z_1(x) - Z_0(x) = \frac{\tilde{\beta}}{\beta} ln \frac{P_1(x)}{P_0(x)} + \left(1 - \frac{\tilde{\beta}}{\beta}\right) [u_1(x) - u_0(x)] \tag{17}$$

The first block consists in equation (17) for all different states $x \in \mathcal{X}$.

To build the second block take equation (16) and apply the normalization to get

$$V(x) = Z_0(x) + ln\left\{exp[Z_1(x) - Z_0(x)] + 1\right\} + (1-\tilde{\beta})\delta \sum_{j=1}^{1} \left[\frac{exp[Z_j(x)]}{exp[Z_0(x) + Z_1(x)]} \sum_{x' \in \mathcal{X}} V(x')\pi(x'|x, j) \right] \tag{18}$$

Note that for each state $x \in \mathcal{X}$ and given$\{\left\langle \delta, \beta, \tilde{\beta} \right\rangle \langle Z_i(x) : i \in \mathcal{I},\, x \in \mathcal{X} \rangle\}$, equation (18) represents a system of $card(\mathcal{X}) = X$ linear equations with X unknowns, namely, $\boldsymbol{V} \equiv [V(1), ..., V(X)]^T$.

It is possible and convenient to express $\boldsymbol{V}$ in matrix notation. For that purpose let $\boldsymbol{A} = Z_0(x) + ln\left\{exp[Z_1(x) - Z_0(x)] + 1\right\}$ be a vector of dimension $X \times 1$; let $\tilde{\boldsymbol{P}} \equiv \left[\tilde{\boldsymbol{P}}_0, \tilde{\boldsymbol{P}}_1\right]$ be the matrix of dimension $X \times 2X$ that contains the choice probabilities; and let $\boldsymbol{\Pi}$ be a matrix of dimension $2X \times X$ that captures the matrices of transition probabilities[14]. Then the system

[14] Where $\boldsymbol{\Pi} \equiv [\boldsymbol{\Pi_0}, \boldsymbol{\Pi_1}]^{\boldsymbol{t}}$ and $\boldsymbol{\Pi_i} = [\boldsymbol{\Pi_i(1)}, ..., \boldsymbol{\Pi_i(X)}]$, in which $\boldsymbol{\Pi_i(x)} = [\pi(1|x, i), ..., \pi(X|x, i)]^t$.

of equations generated by equation (18) can be expressed as:

$$\boldsymbol{V} = \left[\boldsymbol{I} - \left(1 - \tilde{\beta}\right)\delta\tilde{\boldsymbol{P}}\boldsymbol{\Pi}\right]^{-1}\boldsymbol{A} \tag{19}$$

Combining system (19) with equation (10) for all states $x \in \mathcal{X}$ and $i \in \mathcal{I}$ results in the second block that consists of $2 \times X$ equations.

$$Z_i(x) = u_i(x) + \tilde{\beta}\delta\boldsymbol{\Pi_i}(x)\left[\boldsymbol{I} - \left(1 - \tilde{\beta}\right)\delta\tilde{\boldsymbol{P}}\boldsymbol{\Pi}\right]^{-1}\boldsymbol{A} \tag{20}$$

Without loss of generality normalize $u_0(x) = 0$ for all $x \in \mathcal{X}$. Hence the two blocks defined in (17) and (20) yield to $2 \times X$ values for $\{Z_i(x) : i \in \mathcal{I},\, x \in \mathcal{X}\}$ and $1 \times X$ values for $\{u_1(x) \,:\, x \in \mathcal{X}\}$. The number of unknowns is the same as the number of equations in the system, $3 \times X$. Solving this system while taking as given values of $\left\langle \delta, \beta, \tilde{\beta} \right\rangle$ identifies $\langle\{u_i(x),\ Z_i(x'),\ V_i(x') : i \in \mathcal{I},\ x \in \mathcal{X},\ x' \in \mathcal{X}\}\rangle$.

Second Step

In order to identify $\left\langle \delta, \beta, \tilde{\beta} \right\rangle$ the following assumption on the states variables has to be made.

Assumption 5. ***(Exclusion Restriction)***: There exists state variables $x_1 \in \mathcal{X}$ and $x_2 \in \mathcal{X}$ with $x_1 \neq x_2$, such that

1. For all $i \in \mathcal{I}$, $u_i(x_1) = u_i(x_2)$
2. But for at least one $i \in \mathcal{I}$, $\pi(x'|x_1, i) \neq \pi(x'|x_2, i)$

For notation simplicity divide the state variables as follows, (x_r, x_e), where $x_r \in \mathcal{X}_r$ refers to the state variables that affect the instantaneous utility function $u_i(x_r)$, and $x_e \in \mathcal{X}_e$ refers to the state variable that satisfies assumption 5. Instead of presenting the formal argument, I will outline how the use of exclusion variables allows a distinction of δ from β and β from $\tilde{\beta}$.

To see how it is possible to separate β from δ, suppose that a state worker is an exponential discounter with discount factor $\hat{\delta}$. By plugging $\tilde{\beta} = 1$ in equation (18) her expected continuation payoff becomes

$$V(x) = Z_0(x) + ln\left\{exp[Z_1(x) - Z_0(x)] + 1\right\} \tag{21}$$

Which as aforementioned only depends on $ln\frac{P_1(x)}{P_0(x)}$. In other words the expected continuation payoff for an exponential discounter is completely determined by the observed choice probabilities.

Now consider a sophisticated present biased state worker. Her current self values differently the future payoffs than her perceived future self. This incongruence leads to an additional term,

$$(1-\beta)\delta\sum_{j\in\mathcal{I}}\left[\frac{exp[Z_j(x)]}{exp[Z_0(x)+Z_1(x)]}\sum_{x'\in\mathcal{X}}V(x')\pi(x'|x,j)\right]$$

in equation (18). Hence the expected continuation payoff for a sophisticated present biased state worker is more than just observed choice probabilities. As discussed in the first step continuation utilities will determine the identified values of $\{u_1(x) :, x \in \mathcal{X}\}$. By making $u_1(x)$ independent of x_e, it is possible to distinguish β from δ.

β and $\tilde{\beta}$ can be distinguished as follows: Suppose that $\delta = 1$, and note that equation (17) contains $\frac{\tilde{\beta}}{\beta}$. If $\frac{\tilde{\beta}}{\beta} = 1$ then $Z_1(x) - Z_0(x) = ln\frac{P_1(x)}{P_0(x)}$. Hence $\{u_1(x) :\ x \in \mathcal{X}\}$ will be pinned down from the data, which could be refuted if the identified values of $u_1(x)$ do not satisfy the exclusion restriction. Hence $\beta, \tilde{\beta}$ are identified. Proposition 1 in Fang and Wang (2015) provides a formal argument.

B. Estimation

By the identification argument the estimator of $\left\langle\beta, \tilde{\beta}, \delta\right\rangle$ consists in a two step approach. The first step consists of estimating the choice probabilities (for all states) of making an additional contribution. It also requires estimating the transition probabilities $\pi(x'|x,i)$ for all $i \in \mathcal{I}$ and all $(x', x) \in \mathcal{X}^2$.

The second requires solving the following equation system for all $x \in \mathcal{X}$

$$Z_1(x) = u_1(x) + \tilde{\beta}\delta\mathbf{\Pi_1}(x)\left[\boldsymbol{I} - \left(1-\tilde{\beta}\right)\delta\tilde{\boldsymbol{P}}\mathbf{\Pi}\right]^{-1}\boldsymbol{A}$$
$$Z_0(x) = \tilde{\beta}\delta\mathbf{\Pi_0}(x)\left[\boldsymbol{I} - \left(1-\tilde{\beta}\right)\delta\tilde{\boldsymbol{P}}\mathbf{\Pi}\right]^{-1}\boldsymbol{A}$$
$$Z_1(x) - Z_0(x) = \frac{\tilde{\beta}}{\beta}ln\frac{P_1(x)}{P_0(x)} + \left(1-\frac{\tilde{\beta}}{\beta}\right)[u_1(x)]$$

The solution of the system yields values of $Z_1(x)$, $Z_0(x)$ and $u_1 = \hat{u}_1(x_r, x_e)$ for a given triple of $\left\langle\beta, \tilde{\beta}, \delta\right\rangle$ for all $x \in \mathcal{X}$. Such utility values need to satisfy assumption 5, hence it is important to impose the following restriction:

$$\hat{u}_1(x_r) = \frac{1}{|\{(x_r, \tilde{x}_e) : \tilde{x}_e \in \mathcal{X}_e\}|}\sum_{\{(x_r,\tilde{x}_e):\tilde{x}_e\in\mathcal{X}_e\}}\hat{u}_1(x_r, \tilde{x}_e) \tag{22}$$

Given $\hat{u}_1(x_r)$ as defined in equation (22), it is possible to predict the choice probabilities $\hat{P}\left(x; \langle\beta, \beta, \delta\rangle\right)$ and then formulate a pseudo-likelihood from the observed data:

$$\mathcal{L} = \prod_{n\in N}\prod_{i=0}^{1}\prod_{x\in\mathcal{X}}\hat{P}\left(x;\left\langle\beta,\tilde{\beta},\delta\right\rangle\right)^{\Phi_n(x)}\left[1-\hat{P}\left(x;\left\langle\beta,\tilde{\beta},\delta\right\rangle\right)\right]^{1-\Phi_n(x)}$$

Where n stands for an individual and N stands for the number of individuals in the sample, and $\Phi_n(x)$ is a function that takes the value of 1 when the individual makes an extra contribution. I maximize the pseudo-likelihood function to estimate $\left\langle\beta, \tilde{\beta}, \delta\right\rangle$. As shown in Fang and Wang (2015), this estimator is consistent and asymptotically normal, with asymptotic variance given by

$$H = \sum_{x\in\mathcal{X}}Var\left(\frac{\Phi_n(x) - \hat{P}\left(x;\left\langle\beta,\tilde{\beta},\delta\right\rangle\right)}{\hat{P}\left(x;\left\langle\beta,\tilde{\beta},\delta\right\rangle\right)\left(1-\hat{P}\left(x;\left\langle\beta,\tilde{\beta},\delta\right\rangle\right)\right)}\frac{\partial\hat{P}(x)}{\partial\theta}(\theta)\right)$$

therefore standard errors are easily computed.

C. Econometric Specification

Now I proceed to describe how I perform the estimation. First note that a state in the model is defined as a combination of: the outstanding level of debt M_t, the level of income y_t, the gender of the worker, the age when the worker signed the contract and whether she lives in Mexico City. Thus estimating the choice probabilities can be done by a logistic regression through assumption 4. To perform this estimation I assume robust standards errors.

Then, in order to non-parametrically estimate the transition probabilities, I need to discretize the income level and the outstanding level of debt. I divide both variables into 6 equidistant groups, where income level ranges from 0 to 125 or more minimum wages and the outstanding debt goes from 0 to 7500 or more minimum wages. For simplicity purposes I also consider age to be a binary variable that captures if the worker was 40 years old or more when she signed the contract. Therefore I define a state to be a combination of $x = (M, y, Female, 40years, Mexico\,City)$, thus I can observe how many individuals in x move to x'. Hence the estimator of the transition probabilities is

$$\hat{\pi}(x'|x, i) = \frac{n(x'|x)}{n(x)}$$

where $n(\cdot)$ is the counting function. In other words, it is the ratio between the number of workers at state x' that were at x and the total number of workers at state x.

In step two, I propose initial values of $\left\langle \beta, \tilde{\beta}, \delta \right\rangle$, then I solve the system. Solving the system could yield to one solution, multiple solutions, or no solution. If it is the case that there are multiple solutions, I choose the one that obtains the greater pseudo-likelihood value. If the system has no solution, I assign for the likelihood a negative number that is sufficiently large. Once a solution has been found, I impose the restriction (22) unto the utilities. To do so I define my exclusion variables to be $x_e = (Female, 40years, Mexico\,City)$. From the data description section we observed that these variables suggest different assignments of income levels, so they indeed affect the transition probabilities. Since those state variables are time invariant, it is reasonable to assume that they do not affect the difference in instantaneous payoffs of making or not an additional contribution. Once the values of $\hat{u}_i(x_r)$ are obtained, I predict the probabilities under a logistic distribution. This procedure is done recursively until the computer finds an optimum. To ensure global optimality, several initial values are provided, and different optimization methods are used.

V. Results and Policy simulations

As discussed in the previous section the estimation consists of two steps. This section summarizes the results of both steps and presents some comments and interpretations about them. The first step consists of estimating the choice probabilities and the transition probabilities.

A. First Step Results

The choice probabilities are estimated under a logistic regression. I present three different model specifications in which the exclusion variable set changes. All models consider the relevant payoff outcomes in the utility function: outstanding debt and income. Model 1 only includes gender as an exclusion variable; model 2 adds age above 40; model 3 also considers if they live in Mexico City or not. Table 6 summarizes the information on the three specifications and the results for each model. Note that income seems to be positively correlated with making an extra contribution, on the other hand the outstanding debt seems to be negative correlated. Although this relationship is not causal, it seems to be consistent with present biased behavior. As noted before a present biased individual will be more likely to make additional payments later in time, where on average the debt is lower and income is higher. All exclusion variables turn out to be significant, and all of them are negatively correlated with making an additional contribution. Since I am not interested in how much each variable affects the choice probability, but in how well I can predict them instead, the pseudo R^2 is reported. In all specifications the pseudo R^2 takes a value above .4.

The transition probabilities are estimated non-parametrically. In all models the states are defined by a combination of: one of the 6 levels of income, one of the 6 levels of outstanding debt, one combination of the exclusion state variables. Therefore for every $i \in \mathcal{I}$ the dimension of the transition probability matrix when only gender is included (model 1) is 72×72. When age above 40 is included (model 2) this number grows to 144×144. When adding if they reside in Mexico City (model 3) the matrix dimension grows to 288×288.

Figures 5.1 and 5.2 present a subsample of the estimated transition probabilities. In both figures I consider an individual that started with an outstanding debt between 100 and 150 minimum wages, his (her) fortnight income ranged between 2 and 3 minimum wages, when he (she) signed his (her) mortgage he (she) was at most 39 years old, and he (she) lives outside Mexico City. These figures show the next period probability for the individual to move to any other income-debt state. The only difference between figure 5.1 and 5.2, is that figure 5.1 considers a male state worker whereas 5.2 considers a female one. Since I

Table 6—Logit Regression of Choice Probabilities

	(I)	(II)	(III)
Outstanding debt	-0.0033***	-0.0037***	-0.0036***
	(0.0003)	(0.0003)	(0.0003)
Income	1.4300***	1.4932***	1.4852***
	(0.2930)	(0.0306)	(0.0299)
Female	-0.2836***	-0.2353***	-0.2434***
	(0.0488)	(0.0482)	(0.0484)
40 years old	/	-0.7504***	-0.8364***
	/	(0.0534)	(0.0568)
Mexico City	/	/	-0.3795***
	/	/	(0.0566)
Constant	-7.5694***	-7.3553***	-7.1472***
	(0.1251)	(0.1214)	(0.1216)
Pseudo R^2	0.4041	0.4139	0.4162
Observations	110,356	110,356	110,356

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. Income and outstanding debt are measured in minimum wages. Robust standard errors are presented in parenthesis. ***,**,* represent statistical significance at 1%, 5% and 10% respectively.

have a total of 36 income level and outstanding debt combinations, I choose to present them by overlapping the 6 levels of income over the 6 outstanding debt cutoffs. For example in figure 5.1, the worker will most likely move to a state between 50 and 100 minimum wages of outstanding debt, and his income should remain constant. Making an additional payment in both figures increases the probability of lowering the mortgage debt. Also a female state worker has more volatility regarding new income states.

B. *Second Steps Results*

Table 4 presents the estimates for $\left\langle \beta, \tilde{\beta}, \delta \right\rangle$ for all three sets of exclusion variables. The results show that the average worker with a FOVISSSTE mortgage is present biased $\beta \in (0.34, 0.56)$, has a high degree of naivety (in one specification is completely naive $\tilde{\beta} = 1$) and her exponential discount is similar to what the literature has found, $\delta \in (0.85, 0.90)$. Because the estimated value of $\tilde{\beta}$ lies on the parameter space boundary, the correction of Moran (1970) for standard errors is applied. Note that the confidence intervals for this estimation are very tight and all the estimated values are statistically different from zero.

More interesting hypothesis tests are presented in table 5, where it is possible to see if the classical economic theory is refuted. I will focus in the second model since it is the

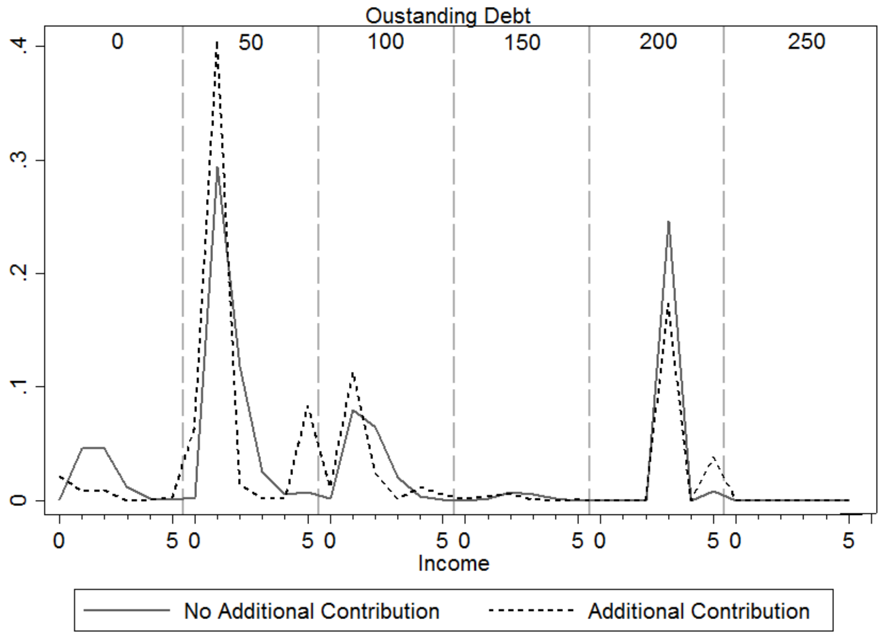

Figure 3. Transition Probabilities for a Male State Worker

Note: Non parametric estimates of the transition probabilities

one that obtains the largest pseudo-likelihood. The first test rejects the absence of present bias, $\beta = 1$. This implies that the average worker will consume as much as possible in the first periods, in fact if the worker exits the public labor force her probability of default should be greater that the one considered under standard models. The second test rejects the possibility of being aware of her own present bias $\beta = \tilde{\beta}$. As aforementioned having an illiquid asset, such as a mortgage, helps smooth consumption. However when an agent ignores her own present bias weak commitment devices are not used, and thus, we observe less additional payments. So, would happen if the worker was not present biased? or How much would her behavior change if she were sophisticated? The next next subsection explores this scenarios.

C. Counterfactual Simulations

Table 6 describes the average predicted probability of making an additional contribution in several scenarios. The first row shows the predicted values for the estimated parameters

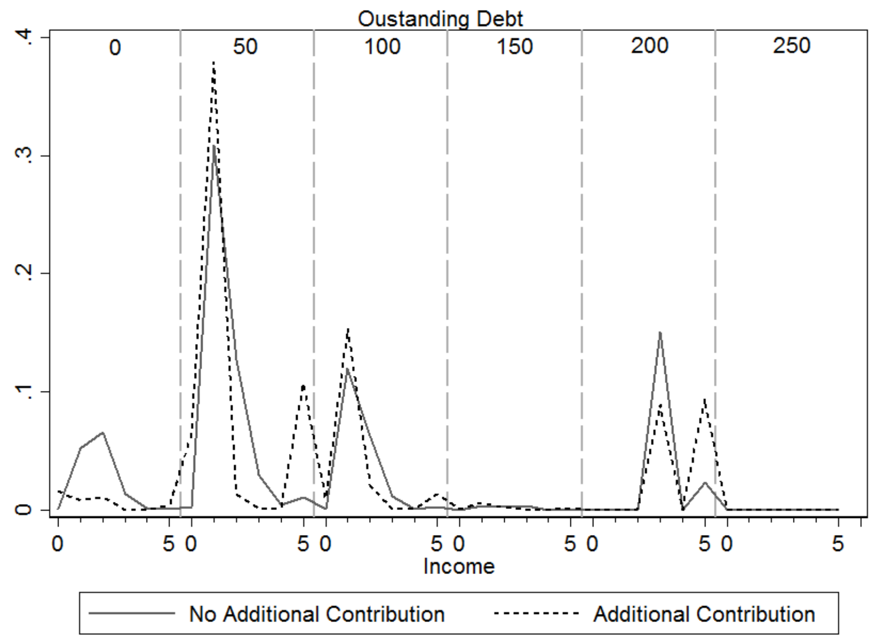

Figure 4. Transition Probability for a Female State Worker

Note: Non parametric estimates of the transition probabilities

in model 2. When comparing them with the case of the worker that discounts exponentially (second row), its possible to see an increase in the probability of making an additional contribution. There is also an increase on the average instantaneous utility of making a payment (across states). The same prediction occurs when we compare it to a sophisticated worker, nevertheless the change are smaller.

So far the economic models assume that individuals cannot transit from having quasi-hyperbolic preferences towards classic exponential ones. Nevertheless there is room for policy making to sophisticate individuals. How to do it escapes the scope of the present paper and further neuroeconomics research is needed. Since a sophisticated worker will have on average a greater probability of making an additional payment, FOVISSSTE could offer a contract that discounts more than 30 percent of their base salary as payment. And so, a better commitment device would be available. Finally, recall that some of the incentives to repay the debt faster come from the fact that the mortgage is measured in minimum wages, thus moving the scheme to one measured in Mexican pesos might lower the default

Table 7—Discount Parameters

Parameter	(I)	(II)	(III)
δ	0.8532**	0.9960***	0.9023***
	(0.0216)	(0.0001)	(0.0132)
	[0.7975 0.9089]	[0.9957 0.9962]	[0.8682 0.9363]
$\tilde{\beta}$	0.8010***	1.000***	0.9998***
	(0.0131)	(0.0759)	(0.0563)
	[0.76720 0.83479]	[0.5236 1.4763]	[0.8548 1.1448]
β	0.5634***	0.3463***	0.4875***
	(0.0010)	(0.0002)	(0.0007)
	[0.56087 0.5695]	[0.3452 0.3473]	[0.4869 0.4893]
LogLikelihood	476.3201	564.3760	527.3210

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. Income and outstanding debt are measured in minimum wages. Robust standard errors are presented in parenthesis. ***,**,* represent statistical significance at 1%, 5% and 10% respectively.

Table 8—Classic Theory Tests/ Under Specification 2

Hypotheses	Decision	Wald-Statistic
Ho: $\beta = 1$	Reject	2455.55
Ho: $\beta = \tilde{\beta}$	Reject	3284.44

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000. The Rejection of the null hypotheses is achieved at 99 percent of confidence.

probability for the naive present biased workers.

Table 9—Simulations of Counterfacutals

Model	Predicted Probability	Instantaneous utility
Estimated	0.03425	3.4222
$\beta = \tilde{\beta} = 1$	0.08030	3.7009
$\tilde{\beta} = \beta$	0.06419	3.4243

Note: Random sample of 410 individuals that started to repay their debt in Jan 2000.

VI. Discussion

In this paper I use novel data from a de-centralized Mexican public mortgage institution to estimate a dynamic structural model with QHD. Using the methodology in Fang and Wang (2015), I am able to identify and estimate the long run discount parameter, the present bias discount parameter and the parameter that captures the state worker's naivety degree.

Identification is achieved by using exclusion variables that affect the transition probability of outstanding debt and income, but not the instantaneous payoffs of making an additional payment to their mortgage. To that end I used gender, age and whether the worker lives in Mexico City as exclusion variables. It is worth mentioning that I am exploiting the long periodicity of the panel to estimate the stationary transition probabilities. In order to relax the stationarity assumption more subjects are needed in the current data set.

Conditional on assuming that there are no other sources of income, I find that the average state worker has a standard long run discount parameter $\delta = 0.9960$, is present biased $\beta = 0.3463$, and does not have perfect awareness of it $\tilde{\beta} = 1.0$. These findings are robust to distinct specifications of the exclusion variables set. I reject the hypotheses of the classical economic theory (exponential discounting), and I also reject that the average individual is sophisticated. I provided counterfactuals that suggest that a policy that educates the state worker will increase the probability of making an additional payment. How to do so remains an open question.

When policy makers ignore the fact that the mortgage market contains present biased individuals, the risk of default is underestimated. Since 2009 FOVISSSTE started to issue debt through the securitization of their mortgages. Under the assumption of time consistent individuals, and since there exist an automatic payment from the state worker, these securities are cataloged as risk free. Still over the last years their nonperforming mortgages have increased from 2 percent in 2002 to 8 percent in 2014 (SHF 2015).

To account for present bias, I suggest that FOVISSSTE should offer a contract that deducts more than 30 percent of their base salary. In this way there exists an option of a stronger commitment device. I also mentioned that lending Mexican pesos instead of minimum wages could reduce the incentives to repay their mortgage faster. This could be beneficial for those who want to repay faster, but who are also dissuaded from doing so due to their present bias. In 2015 FOVISSSTE announced their first credit scheme in Mexican pesos; however, it is too soon to see if that measure has been beneficial, and thus it opens an opportunity for research.

From a behavioral economics perspective this paper contributes to augment the literature on present bias, moreover it contributes to the short list of non experimental structural estimation on behavioral economics.

REFERENCES

Andreoni, James, and Charles Sprenger. 2012. “Estimating time preferences from convex budgets.” *The American Economic Review*, 102(7): 3333–3356.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. 2001. “The hyperbolic consumption model: Calibration, simulation, and empirical evaluation.” *The Journal of Economic Perspectives*, 15(3): 47–68.

Arceo-Gomez, Eva O, and Raymundo M Campos-Vazquez. 2014. “Race and marriage in the labor market: A discrimination correspondence study in a developing country.” *The American Economic Review*, 104(5): 376–380.

Ariely, Dan, and Klaus Wertenbroch. 2002. “Procrastination, deadlines, and performance: Self-control by precommitment.” *Psychological science*, 13(3): 219–224.

Bryan, Gharad, Dean Karlan, and Scott Nelson. 2010. “Commitment devices.” *Annu. Rev. Econ.*, 2(1): 671–698.

De Clippel, Geoffroy, and Kareen Rozen. 2014. “Bounded rationality and limited datasets.”

DellaVigna, Stefano. 2009. “Psychology and economics: Evidence from the field.” *Journal of Economic Literature*, 47(2): 315–372.

DellaVigna, Stefano, and M Daniele Paserman. 2005. “Job search and impatience.” *Journal of Labor Economics*, 23(3): 527–588.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. “Paying not to go to the gym.” *The American Economic Review*, 96(3): 694–719.

Dupas, Pascaline, and Jonathan Robinsona. 2013. “Why don’t the poor save more? Evidence from health savings experiments.” *The American Economic Review*, 103(4): 1138–1171.

Echenique, Federico, Taisuke Imai, and Kota Saito. 2017. “Testable Implications of Models of Intertemporal Choice: Exponential Discounting and Its

Generalizations.”

Fang, Hanming, and Dan Silverman. 2009. “Time-inconsistency and welfare program participation: Evidence from the NLSY.” *International Economic Review*, 50(4): 1043–1077.

Fang, Hanming, and Yang Wang. 2015. “Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions.” *International Economic Review*, 56(2): 565–596.

Hotz, V Joseph, and Robert A Miller. 1993. “Conditional choice probabilities and the estimation of dynamic models.” *The Review of Economic Studies*, 60(3): 497–529.

Jones, Damon, and Aprajit Mahajan. 2015. “Time-inconsistency and saving: Experimental evidence from low-income tax filers.” National Bureau of Economic Research.

Kuchler, Theresa. 2015. “Sticking to Your Plan: Empirical Evidence on the Role of Present Bias for Credit Card Paydown.”

Laibson, David. 1997. “Golden eggs and hyperbolic discounting.” *The Quarterly Journal of Economics*, 112(2): 443–478.

Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2017. “Estimating discount functions with consumption choices over the lifecycle.” National Bureau of Economic Research.

Magnac, Thierry, and David Thesmar. 2002. “Identifying dynamic discrete decision processes.” *Econometrica*, 70(2): 801–816.

Meier, Stephan, and Charles Sprenger. 2010. “Present-biased preferences and credit card borrowing.” *American Economic Journal: Applied Economics*, 2(1): 193–210.

Moran, Patrick AP. 1971. “Maximum-likelihood estimation in non-standard conditions.” Vol. 70, 441–450, Cambridge University Press.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing it now or later." *American Economic Review*, 103–124.

Oshry, Barry. 2006. "Advancing beyond advances in behavioral economics." *Journal of Economic Literature*, 44(3): 694–711.

Paserman, M Daniele. 2008. "Job search and hyperbolic discounting: Structural estimation and policy evaluation." *The Economic Journal*, 118(531): 1418–1452.

Phelps, Edmund S, and Robert A Pollak. 1968. "On second-best national saving and game-equilibrium growth." *The Review of Economic Studies*, 35(2): 185–199.

Robson, Arthur J, and Larry Samuelson. 2009. "The evolution of time preference with aggregate uncertainty." *The American Economic Review*, 99(5): 1925–1953.

Rubinstein, Ariel. 2003. ""Economics and psychology"? The case of hyperbolic discounting." *International Economic Review*, 44(4): 1207–1216.

Rust, John. 1994. "Structural estimation of Markov decision processes." *Handbook of econometrics*, 4: 3081–3143.

Strotz, Robert Henry. 1955. "Myopia and inconsistency in dynamic utility maximization." *The Review of Economic Studies*, 23(3): 165–180.

Thaler, Richard H, and Shlomo Benartzi. 2004. "Save more tomorrow: Using behavioral economics to increase employee saving." *Journal of political Economy*, 112: S164–S187.

Topelson, Sara. 2014. "Estado Actual de la Vidienda en Mexico." *Sociedad Hipotecaria Federal.*