RETIREMENT IN THE SHADOW (BANKING)

Guillermo Ordoñez
Facundo Piguillem

Retirement in the Shadow (Banking)
Guillermo Ordoñez and Facundo Piguillem
NBER Working Paper No. 26337
October 2019
JEL No. E21,E44,G21,J11

ABSTRACT

The U.S. economy has recently experienced two, seemingly unrelated, phenomena: a large increase in post-retirement life expectancy and a major expansion in securitization and shadow banking activities. We argue they are intimately related. Agents rely on financial intermediaries to save for post-retirement consumption. When expecting to live longer, they rely more heavily on intermediaries that use securitization, with riskier but higher returns. A quantitative evaluation of the model shows the potential of the demographic transition to account for a boom in credit and output, but only when it triggers a more extensive use of securitization and shadow banking.

Guillermo Ordoñez
University of Pennsylvania
Department of Economics
428 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
ordonez@econ.upenn.edu

Facundo Piguillem
EIEF
Via Sallustiana, 62 00187
Roma. Italia
facundo.piguillem@eief.it

1 Introduction

In the last four decades the U.S. economy experienced a steep increase in intermediated credit, which almost doubled from one to two GDPs. Due to the magnitude of the 2008 financial crisis, policymakers and scholars rationalized this "credit boom" in different ways, ranging from an atypical influx of foreign funds (an international savings glut) to pure financial speculation. By focusing, perhaps excessively, on its role to trigger crises, these explanations tend to deny the boom's potential benefits and deem it just as a detrimental phenomenon. But was there any gain from the credit expansion? If so, how large were these gains?

We analyze the contribution of a domestic factor that has been under-emphasized to explain this prolonged credit boom: the demographic transition characterized by a longer life span. In just four decades, life expectancy of the U.S. population conditional on retirement, increased dramatically from 77 years to around 83 years. Even though life expectancy has been increasing for a century, this time frame was unique in that it was driven by people getting older as opposed to previous decades in which it was driven by a dramatic decline in child mortality.[1]

Living longer after retirement induces an increase in the demand of savings, which we argue opens the doors for new and more efficient tools and institutions to supply savings – *securitization* and the so-called *shadow banking*. This view is in line with the recent work of Scharfstein (2018) who, using cross-country evidence, highlights that pension policies and other *restrictions to save for retirement* affect the structure of financial systems, in particular the balance between banks and capital markets. In a similar vein we argue, using time series evidence for the U.S., that the *needs to save for retirement* affect the type of financial instruments and the composition of banks that are used.[2]

Savings for retirement needs are indeed considered of enormous relevance to study macroeconomics, as retirees hold a large fraction of total wealth in the economy. Wolff (2004) documents that more than a third of total wealth in the United States is held by households whose heads are over 65, and Gustman and Steinmeier (1999) show

---

[1]The average retirement age in the U.S. is 63.5 years. For the historical evolution of life expectancy, see https://www.cdc.gov/nchs/data/hus/2011/022.pdf.

[2] The current debate about the falling "savings rate" in the U.S. economy may seem to contradict the higher pressure of savings for retirement needs. As we explain in detail in Section 4.1, however, our theory is consistent with the observed fall in the saving rate *out of disposable income*.

that for households near retirement, wealth is around one-third of lifetime income. Even before retirement, Gale and Scholz (1994) and Kotlikoff and Summers (1981) argue that most people's savings are intended to be used after retirement. It is not surprising then the existence of a rich literature on the *macroeconomic effects of pensions*, dating as back as the celebrated overlapping generations model of Samuelson (1958).

What is puzzling, however, is the little existing work on the *financial markets effects of pensions*, given that in the U.S. a large fraction of such financial wealth is managed by financial intermediaries. The first panel of Figure 1 shows that intermediaries manage at least 50% of total wealth (accounts in pension funds, deposits in banks and shares in mutual funds), with the composition moving away from traditional deposits and towards mutual funds shares, both directly (the black line in the first panel) and indirectly through changes in the investments of traditional intermediaries (such as the increase of mutual fund shares held by private pension funds, the black line of the second panel of Figure 1).[3]

Figure 1: Evolution of Shadow Banking
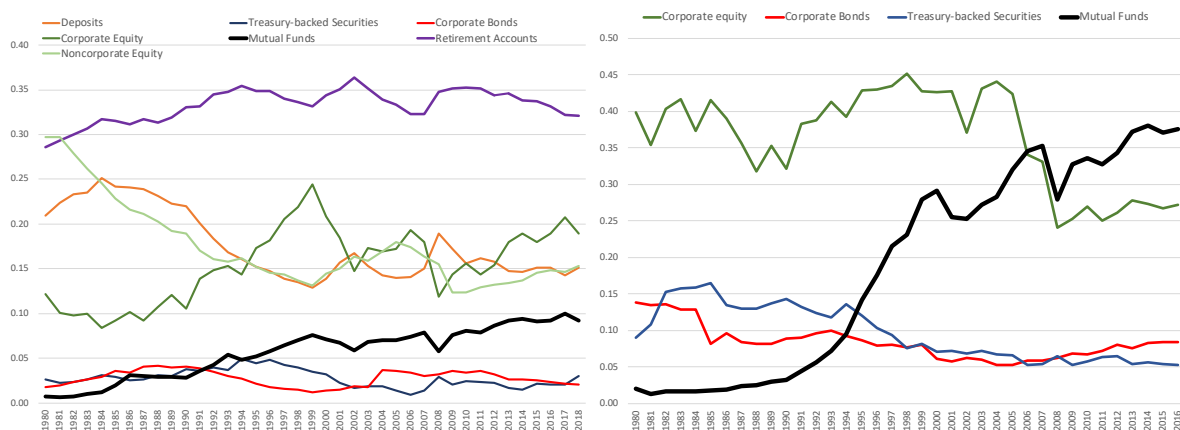
In this paper we connect *theoretically* the increase in life expectancy that raises savings for retirement needs with the increase in shadow banking (through the extensive use of securitization) that characterized the changes in the anatomy of financial intermediation. We then evaluate this connection *quantitatively*. In particular, we show that *i)* shadow banking was instrumental in accommodating the larger saving needs, and

---

[3]This aggregate evidence is also consistent with the more detailed portfolio composition based on the Vanguard Research Initiative reported by Ameriks et al. (2014).

it did so by substantially resorting to securitization and decreasing the financial sector's liquidity cost; *ii)* this *domestic savings glut* can account for most of the observed credit boom; and *iii)* even if we assume that the great recession was entirely caused by shadow banking operations, the benefits prior to the crisis were an order of magnitude larger than the cost of the crisis.

To study the macroeconomic implications of these demographic and financial developments, we proceed in four stages. *The first stage is theoretical* (Section 2). We propose an overlapping generations model with heterogeneity in the bequest motives that allows for the coexistence of lenders and borrowers. Individuals with low-bequest motives save for retirement by depositing their funds in *financial intermediaries (or banks)*. Individuals with high-bequest motives save for retirement by buying stocks in *capital markets*.[4] While Scharfstein (2018) focuses on this last margin, we instead explore the composition of bank types.

In our economy banks perform two costly activities, *i)* they channel credit from depositors (low-bequest motives) to investors (high-bequest motives) and *ii)* as deposits are short-term contracts that can be withdrawn at any period, banks have to guarantee their availability upon withdrawal. We justify holding short-term deposits as being disciplining devices for banks, as in Diamond and Rajan (2001), which imposes having enough resources to prevent bank runs as in Diamond and Dybvig (1983). The cost of providing credit, which we denote *operation cost*, is the cost of finding the best available investment opportunities to allocate funds, including the process of identifying productive opportunities, monitoring the management of projects and administering payments. The cost of guaranteeing the availability of deposits, which we denote *liquidity cost*, is the cost of transforming long-term risky loans into short-term safe assets that can be liquidated at stable nominal conditions in relatively short periods of time in case a large fraction of depositors decide to withdraw their funds.

There are two types of financial intermediaries: *traditional banks* and *shadow banks*. In contrast to traditional banks, shadow banks can use securitization to reduce liquidity costs. Securitization is a technology that involves transforming a pool of assets into a new financial instrument (*security*) that improves the liquidity in the marketplace of the assets being securitized. This technology of pooling assets, dividing the pool

---

[4]The relevance of bequest motives for retirement savings has been discussed by Bernheim (1991) and Lockwood (2012), among others. Alternatively, however, one could think about individuals with high bequest motives as being risk takers or having access to productive investments and agents with low bequest motives as safety seekers.

into tranches and making transactions complex and opaque discourages asymmetric information among market participants, facilitating trading and improving the liquidity of underlying assets. By operating at lower liquidity costs, shadow banks can offer better rates for deposits, but at a cost in terms of fragility (sudden dry up of liquidity) inherent to the use of opaque operations. Gorton and Ordonez (2014) provide a model of this trade-off. Ceteris paribus, a higher life expectancy triggers an appetite for yields, and depositing in shadow banks become more appealing to fulfill that appetite.

But how relevant were shadow banking operations in the United States to reduce liquidity costs? To address this question, *the second stage is empirical* (Section 3). Measuring the quantitative extents and implications of securitization is challenging because of its ubiquitous use in financial markets, its lack of transparency and the corresponding issues of double counting. Our approach is to use prices instead of quantities. We use the model to map the use of securitization into a "liquidity premium" that can be inferred from measuring the spread between lending and deposit rates in the whole financial sector. Measured this way, the liquidity cost declined from a stable level of 1% in 1980 to almost 0% before the recent financial crisis. This finding is consistent with two alternative estimates using unrelated methodologies. First, Corbae and D'Erasmo (2018) show, using individual bank's data from FDIC Call Reports, that regulatory reported spreads have declined by around 1% during the same timeframe. Concurrently, Del Negro et al. (2017b) show, using time series for interest rates, that the natural U.S. interest rates have been continuously declining since the 1980's and that 1% of the decline is explained by the increase in the convenience yield.

Is this estimated decline in liquidity costs quantitatively consistent with the changes in volumes and prices of intermediated credit observed in the U.S. since 1980? What were the individual contributions of higher retirement needs and of shadow banking operations to growth and output? To answer these questions, *the third stage is quantitative* (Section 4). We calibrate the economy to 1980 and input the change in life expectancy and liquidity costs to generate a counterfactual for 2007.

Only including these two forces we can account for the observed evolution of households' debt over GDP and total financial assets held in the economy, with an increase of around 75% in both figures by 2007. On the one hand, absent shadow banking, the change in life expectancy would not be able to account for any increase in household debt over GDP, but just a steep decline in the risk-free rate. On the other hand, ab-

sent demographic changes the risk-free rate would have increased substantially and steady state output would have grown by only half as with both forces combined. These results highlight the importance of first understanding the determinants of financial markets to then assessing their impact on aggregate dynamics.

Why securitization, by reducing liquidity costs, affects aggregate real variables in this way? First, securitization improves the liquidity of productive risky loans, allowing banks to invest more on them and less on unproductive government bonds. Second, shadow banking allows banks to escape blunt, and potentially restrictive, regulatory constraints that inefficiently impose the condition that a fraction of assets be invested in unproductive asset classes. These two forces allow for banks to channel resources more efficiently to the most productive opportunities, as in Ordonez (2018a).

Our model abstracts from the possibility that shadow banking collapses on path. Yet, we discuss our initial question, one that has attracted fierce debate in policy and regulatory circles: did the U.S. win or lose from the operation of shadow banks? This justifies our *last, counterfactual, stage* (Section 5). We construct a hypothetical economy without shadow banks and compare it with the realized economy in the U.S. We find that, from 1980 to 2007, the operation of shadow banking increased output by an accumulated 60% of 2007 GDP. This number can be put in context when compared to the cost of the great recession, which we compute to have been of a magnitude of just 14% of 2007 GDP.[5] Thus, even in the extreme case of blaming the crisis and its cost entirely on securitization and shadow banking, still the economy gained (net of the crisis) almost half of 2007 GDP by the operation of shadow banking since the 1980s.

Related Literature: We contribute to a literature that quantitatively studies the determinants of aggregate savings. De Nardi, French, and Jones (2009), De Nardi, French, and Jones (2010), De Nardi, French, and Jones (2015) and Imrohoroglu and Zhao (2018) show that several factors related to *aging*, e.g. health care risks, are relevant drivers of savings. Also, Imrohoroglu, Imrohoroglu, and Joines (1998) point to the sizable impact that assets' *returns* on determining the volume of savings in the economy. Our contribution to this literature is to link aggregate savings driven by aging to the set of financial alternatives, with endogenous heterogeneous returns, that agents

---

[5]To do this calculations we follow the literature and compute the cost by comparing the realized GDP against a potential GDP constructed by the U.S. Congressional Budget Office. The original estimate, before adjusting downwards the potential GDP, generated a cost of the crisis of around 23% of 2007 GDP.

can use to save. More recently, Auclert, Martenet, and Malmberg (2019) study how in a world with aging population the larger needs for retirement insurance could be generating the global decline in interest rates.

We also contribute to the recent academic and policy debate on the effects of shadow banking for macroeconomic aggregates. While most of this debate focuses on the costs of shadow banking in terms of inducing financial fragility, much less is known about its potential positive macroeconomic effects. Moreira and Savov (2015) highlight that shadow banking improves liquidity provision during periods of low economic uncertainty, but focus on the implied fragility and its collapse when uncertainty increases. Begenau and Landvoigt (2017) provide a quantitative model of optimal regulation of traditional banking that recognizes that it may induce the creation of fragile shadow banks. Similarly, Farhi and Tirole (2017) argue that traditional banking is sustained on complementarities between costly public supervision and beneficial public liquidity guarantees, and discuss how regulation (taxes and subsidies, ring fencing, etc.) can accommodate these forces to avoid a migration towards shadow banking.[6]

In this paper we take a longer-term perspective and study the role of demographic changes in boosting the use of shadow banks to better accommodate larger saving needs, focusing on their implications for economic growth and not on their implications for economic cycles nor on how to regulate their activity. Still, even though we focus on the growth and not on the demise of shadow banking, we are able to provide an estimate of its net gains when attributing the whole responsibility of the crisis to its operation.

Our paper is also related to the discussion about the origins of the pressures for channeling more savings. In contrast to a rich literature that argues that the higher demand for safe assets in recent decades can be attributed to larger saving needs of foreign countries (the "global savings glut" hypothesis, as in Caballero (2010), Justiniano, Primiceri, and Tambalotti (2013), Caballero, Farhi, and Gourinchas (2016), Carvalho, Ferrero, and Nechio (2016)), or to larger saving needs of corporations (the "corporate savings glut" hypothesis, as in Bates, Kahle, and Stulz (2009) and Gao, Whited, and Zhang (2018)), in this paper we focus on larger saving needs of longer-living U.S. residents (a "domestic savings glut" hypothesis). Interestingly, a large part

---

[6]Harris, Opp, and Opp (2014) and Plantin (2015) also study the interactions between regulation of traditional banking and shadow banking.

of the savings glut coming from foreign countries has been accommodated by an increase in U.S. government debt and the provision of U.S. government bonds. Shadow banking, then, has had a primary role in accommodating the domestic demand for safe assets, and indeed, we find that these forces are substantial quantitatively and can account for most of observed macroeconomic changes.

The paper also contributes more generally to the discussion on capturing quantitatively the role of banking in macroeconomics using general equilibrium models, as in Diaz-Gimenez et al. (1992) and Mehra, Piguillem, and Prescott (2011). We extend these environments by studying the role of retirement savings in shaping financial markets, in particular the use of securitization and the presence of shadow banking, and in affecting macroeconomic variables not only directly but also indirectly through a new financial system.

To write a parsimonious model suitable to perform a macro quantitative analysis, we have refrained from modeling the microfoundations of shadow banks and how their intensive use of securitization reduces liquidity costs in the system, as in Ordonez (2018b). Instead we rely on reduced-form specifications that are better suited to discipline the model quantitatively using aggregate financial and macroeconomic data.

2   Model

2.1   Environment

We study an overlapping generations economy populated by agents who work in a competitive production sector, save for retirement (either in capital markets or through financial intermediaries) and are taxed by the government.

2.1.1   Agents

Each period a measure $(1 + \eta)^t$ of agents are born, where η is the population growth rate. Agents are born at age $j = 0$ and live with certainty for T periods, during which they work an inelastic amount of hours (normalized to 1) at no utility cost. After age T they can no longer supply labor (they retire) and die with constant probability $0 < \delta < 1$ thereafter. When an agent dies at age j she may leave bequests b_j to her

offspring (a younger agent), which provides a utility $\alpha \geq 0$ (in units of consumption) per unit of bequest. Agents are *only* heterogeneous in the intensity of their bequest motive, $\alpha \sim m(\alpha)$.[7]

We denote the consumption of an age-j agent at calendar time t by $c_{t,j}$ and the discount factor by β. Assuming logarithmic preferences, the utility in present value of an agent who is born at a calendar period t is

$$\sum_{j=0}^{T} \beta^j \log c_{t+j,j} + \sum_{j=T+1}^{\infty} \beta^j (1-\delta)^{j-T-1} [(1-\delta) \log c_{t+j,j} + \delta \alpha \log b_{t+j,j}]. \tag{1}$$

In this specification we make two simplifying assumptions, which are useful for expositional reasons and not overly restrictive. First, we assume just a "joy-of-giving" type of bequest motive, but α may capture other forces as well, such as precautionary savings against possible health shocks.[8] Second, retirement is exogenous at age T. As Costa (1998) and Bloom et al. (2007) show, the retirement age in the U.S., as in many other countries, has been continuously decreasing over the last century. Hence, our assumption is conservative on capturing the effect of aging on savings.[9]

Individuals have three sources of income. First, an agent born in period t receives labor income $y_{t+j,j}$ for the labor provided at age j during her working age. Second, we assume that the bequest $b_{t+j,j}$ that the agent leaves upon death at age j is equally

---

[7]As in Mehra, Piguillem, and Prescott (2011) model of financial intermediation, we rely on heterogeneous bequest motives to capture simultaneously returns differentials and portfolio choices. As is well known from the literature on the equity premium puzzle, Mehra and Prescott (1985), and on the participation puzzle, Mankiw and Zeldes (1991), this is a daunting task using just risk profiles.

[8]As De Nardi, French, and Jones (2010) show, one important motive to save after retirement is to insure against medical expenses. De Nardi, French, and Jones (2015) and Lockwood (2015) point out, however, the difficulty to properly disentangle bequests motives from health needs. Further, our specification is useful to capture a non-trivial age structure of savings. If we had assumed, for instance, that agents are perfectly altruistic with respect to their offspring ("Barro-Becker" type of bequest motive), individual savings would be independent of both life span and survival probabilities, at odds with the empirical evidence discussed by De Nardi, French, and Jones (2009).

[9]As Bloom, Canning, and Moore (2014) argue, as life expectancy increases there are two effects affecting the retirement decision. On the one hand, workers can extend their working life to compensate the longer retirement, but on the other hand, the increase in labor productivity that usually accompanies a longer life increases the demand for leisure (income-wealth effect), which induces an earlier retirement. The net effect of living longer on the retirement age is then ambiguous. Recent work, such as Shourideh and Troshkin (2017), however, point to the dominance of the income-wealth effect, except for individuals in top income decile. Alternative explanations of why agents do not retire older range from an increased female labor force participation, as proposed by Borella, De Nardi, and Yang (2017), to survey-based evidence of demand-side factors such as low expectations of finding flexible jobs, as proposed by Ameriks et al. (2019).

distributed among all agents alive of age $T_I < T$. Thus, every agent receives an inheritance, $\bar{b}_{t+T_I}$, at age T_I. Finally, the agent receives pension transfers P^i_{t+j} from Social Security every period after retirement.[10]

Notice that these three sources of income are deterministic and identical to all agents, so we abstract from aggregate risk and other sources of idiosyncratic risk (such as unemployment or health shocks during the working lifetime) and from heterogeneity in income. This implies that the only source of risk in the economy is the agent's life span and the only saving motive is retirement. This assumption allows us to focus on to the role of aging on financial intermediation, and it is not likely very restrictive for our purposes. First, in spite of underestimating the level of precautionary savings (even though Gale and Scholz (1994) and Kotlikoff and Summers (1981) show that between 75% and 90% of individual savings can be explained by retirement reasons alone), we are interested in their change. Second, although we abstract from the risk premia embedded in interest rates, we are interested on the intermediation spread, not in the level of the interest rate.[11]

In terms of asset accumulation, we will allow agents to differ on their saving strategies depending on their bequest motives. Denoting agent i's saving returns by r^i_t and assuming a labor income tax τ, the agent i born at t has a consolidated total wealth at birth of [12]

$$v^i_t = \sum_{j=0}^{T-1} \frac{(1-\tau)y_{t+j,j}}{\prod_{l=0}^{j}(1+r^i_{t+l})} + \frac{\bar{b}}{\prod_{l=0}^{T_I}(1+r^i_{t+l})} + \sum_{j=T}^{\infty} \frac{(1-\delta)^{j-T}P^i_{t+j}}{\prod_{l=0}^{j}(1+r^i_{t+l})}. \tag{2}$$

We restrict agents' choices to save for retirement to two alternatives: i) *banks* or ii) *capital markets*. Since the only source of uncertainty is the age of death, we assume that households choose among these alternatives at birth and cannot switch across savings alternatives during their lifetime. We take this assumption to mean that there

---

[10]The introduction of social security payments is important because it attenuates the needs for private insurance. Without it, we could be overstating the impact of demographic changes.

[11]When we calibrate the model in Section 4.1, however, we discuss how we adjust interest rates for risk premia to be consistent with this abstraction.

[12]Later we will focus on the balanced growth path. In that case equation (2) greatly simplifies to:

$$v^i_0 = \sum_{j=0}^{T-1} \frac{(1-\tau)y_j}{(1+r^i)^j} + \frac{\bar{b}}{(1+r^i)^{T_I}} + \frac{(1+r^i)}{r^i+\delta} \frac{P^i}{(1+r^i)^T}.$$

is a cost to choose and to switch saving strategies, which has empirical support.[13] To be more precise these two strategies are:

1) **Save in Capital Markets (Strategy C):** Buy equity or bonds in capital markets (corporate equity) or buy and manage an own firm (non-corporate equity) while working and live out savings after retirement, bequeathing any un-spent savings.

2) **Save in Banks (Strategy B):** Sign a contract with a financial intermediary (a bank) that specifies the payment that *the agent must make to the intermediary* during the agent's working age (how much to deposit in the bank or to contribute to the pension plan) and the payment that *the intermediary must make to the agent* when the agent retires. That is, the agent consumes c_j as long as the agent is alive and leaves b_j to her heirs contingent on dying at age j.

We will discuss in detail later how banks operate. It is enough to highlight at this point that there are two differences with capital markets. First, banks constitute a "pool of agents' funds." In this paper this is useful to cross-insure agents who die old with agents who die young, but in general a pool also allows to insure alternative sources of risk. Second, banks manage assets at a cost.

The agent can choose to sign this contract with one of two possible banks: a traditional bank (TB) or a shadow bank (SB). Shadow banks have access to securitization (as we will discuss at length later, securitization reduces the liquidity premium and allows these banks to pay higher rates to their depositors) and signing the contract with a shadow bank implies an additional utility cost κ to the agents. This parameter captures several costs related to securitization, such as the effort cost to understand securities or the uncertainty of participating on a more fragile banking scheme. We model these costs in reduced form here, but in the Appendix we show how they arise endogenously when securitization may collapse, as it happened during the recent U.S. financial crisis.

---

[13]According to the IRS, individuals withdrawing from a retirement plan before 59.5 years old must pay income tax plus an additional 10% on the amount taken out. Consistent with this cost, Mankiw and Zeldes (1991) show evidence that most households don't ever hold stocks and prefer to keep all their assets in risk-less alternatives (participation puzzle). Even households that hold stocks in their portfolios don't drastically change strategies as they age. Fagereng, Gottlieb, and Guiso (2017) argue that a combination of participation costs and a small "disaster" probability are needed to rationalize the low change in investments. Alvarez, Guiso, and Lippi (2012) show that not only are participation costs needed, but also observational costs.

2.1.2 Productive Sector

The productive sector operates every calendar period t with a Cobb-Douglas production function with exogenous growth rate γ,[14]

$$Y_t = K_t^{\theta}(\Gamma_t L_t)^{1-\theta}$$
$$\Gamma_{t+1} = (1+\gamma)\Gamma_t,$$

where K is the aggregate stock of capital in the economy, L is the aggregate supply of labor, Γ is the average labor productivity and θ is the share of capital income over total income. As we discussed in more detail in Section 4.1, we interpreted K not only as productive capital, but also as any kind of storable good, i.e., it includes housing and land. We do this because the wealth to GDP ratio is a key target moment in our calibration.

Labor and capital markets are competitive, which implies that the rental rate of the inputs equals their respective marginal productivity. This is,

$$\delta_k + r_e = F_K(K_t, \Gamma_t L_t)$$
$$y_t = F_L(K_t, \Gamma_t L_t)$$

where δ_k is the capital depreciation rate.

Notice that Γ is labor-augmenting productivity. Thus, because average productivity grows at the rate γ per year, individual wages also grow at rate γ as the agents age: $y_{t+1,j+1} = (1+\gamma)y_{t,j}$.

2.1.3 Government

The government consumes a constant proportion g of output (not valued by agents), follows a committed debt policy D_t^G (independent of prices and quantities in the economy) and pays average Social Security transfer of $\overline{P}_t$. The government collects taxes

---

[14]We are abstracting from changes on the growth rate of productivity, γ. As Chen, Imrohoroglu, and Imrohoroglu (2006) and Fernandez, Imrohoroglu, and Tamayo (2018) show, these changes can have important effects on savings rates. We are studying, however, a time interval in which the U.S. economy can be considered stationary with minor variations in the growth rate of GDP. Nevertheless, in Section 4.3 we incorporate observe changes in productivity, which slightly improves our results.

on labor income to balance the budget,

$$\tau y_t L_t + (D_{t+1}^G - D_t^G) = gY_t + \overline{P}_t + r_{t,L}D_t^G. \tag{3}$$

We will assume hereafter that the Social Security transfer after retirement is a fraction ss^i of the last wage $y_{t,T}$ at retirement, which may be conditional on the saving decisions of individuals. As the only source of heterogeneity in wealth is agents' saving decisions, $i \in \{B, C\}$. Then,

$$P_{t+j}^i = ss^i y_{t+T,T}; \quad \forall j > T.$$

2.1.4 Financial Intermediation

The financial sector consists of perfectly competitive banks that offer saving contracts, specifying the gross rate $1 + r_t$ that an agent receives per unit of saving. With these funds, the bank can invest either in "safe government bonds" that pay with certainty a unit gross rate $1 + r_{L,t}$ per unit of bond or in a continuum of "risky loans" that pay a unit gross rate $1 + \widehat{r}_{e,t} > 1 + r_{L,t}$ per unit of loan. As the bank invests in a continuum of loans, a known fraction s_b of loans default, so there is no ex-ante uncertainty on their return. Each bank takes the return of bonds (that is, $r_{L,t}$) and the risk-adjusted return on loans (that is, $r_{e,t} \equiv (1 - s_b)(1 + \widehat{r}_{e,t}) - 1$) as given. We denote the fraction invested in loans as f_t.

We denote the total financial intermediary's liabilities in period t by D_t. We assume bank's contracts with depositors are short-term, which means D_t should be rolled over every period. There are different ways to justify the empirical regularity that deposits are short-term contracts. Short-term liabilities can have liquidity benefits if agents should use them for transaction purposes (deposits as money, as in Dang et al. (2017)) or to face liquidity shocks (deposits as insurance, as in Diamond and Dybvig (1983)). Short-term liabilities can also induce bankers' discipline, as in Diamond and Rajan (2001). We denote the total financial intermediary's assets in period t by A_t, and assume they mature every period. As banks' liabilities and assets last for a single period, we effectively have a bank's static problem every period and then we dispense from using the calendar subscript t henceforth in this section.

The problem for banks of holding short-term liabilities is that they face the proba-

bility of *bank runs*: all depositors choose to withdraw their funds and move them to another bank. If a bank does not have enough funds to cover these withdrawals, it must default completely on all depositors, which creates a coordination problem for depositors and introduce runs on the equilibrium path. How easily can a bank liquidate its assets on short notice and raise funds to be insulated from this possible coordination failure? The intermediary could raise funds from selling bonds, at a price $1 + r_L$, and from selling its self-originated loans, potentially at a fire-sale price that we denote $1 + q$.

The fire-sale price $1 + q$ that the intermediary can obtain from selling its loans in case of distress depends, however, on how valuable those loans are for potential buyers (other banks not facing a run at the same time). There are many reasons why buyers cannot reap all the benefits of non-originated loans, which range from asymmetric information about the quality of the loans to relationship lending that makes loans more easily monitored by the originator. For a given rate r promised to depositors, the bank is *resilient* (not subject to a bank run) when

$$[z(1+q) + (1-f)(1+r_L)]A \geq (1+r)D, \tag{4}$$

where $z \leq f$ is the amount of loans that are liquidated to face the run.

In terms of the banking technology and market structure, we assume that banks face a constant returns to scale technology, with a constant marginal cost of operation $\widehat{\phi}$ per unit of asset managed, and that there is perfect competition, such that a bank's zero profit condition is

$$[f(1+r_e) + (1-f)(1+r_L) - \widehat{\phi}]A = (1+r)D. \tag{5}$$

Finally, we introduce the next two natural parametric assumptions.

Assumption 1 *There is no arbitrage (agents can buy bonds at no cost). This guarantees*

$$r = r_L.$$

Assumption 2 *Operational costs are not high ($r_e > \widehat{\phi}$). This guarantees*

$$A = D.$$

13

The Role of Securitization: Now we introduce the market for fire sales that determines q, and highlight the role of securitization.

We assume that a bank facing a run (*in distress*) randomly matches with another bank to sell its loans. Since the buyer may not have the expertise to operate the loans, it will try to sell those loans to another bank that is better suited to operate them, obtaining the corresponding return r_e. The probability the buyer can resell those loans is

$$Pr(reselling) = (1 + \Psi) \ln \zeta (1 + z) \frac{1 + r}{1 + r_e}.$$

If the buyer does not find another intermediary willing to buy the loans obtained from the bank in distress, the buyer does not obtain any return.

The reselling probability captures the simplicity of exchanging assets in financial markets. We assume this probability increases in an exogenous parameter $\Psi \geq 0$ that captures the technology available for finding counterparties and for reducing frictions for trading and re-trading assets in the market. As securitization improves trading in secondary markets, relaxing asymmetric information considerations, we model a better securitization technology with a higher Ψ. In order to obtain q in a simple analytical form, we also assume the probability of reselling increases in the amount of acquired loans (because of better pooling possibilities, for instance), decreases in the ratio $\frac{1+r_e}{1+r}$ (a measure of loan specialization vis-a-vis government bonds and other standard assets) and increases in a parameter ζ that we just introduced to guarantee the probability is bounded between 0 and 1 for the relevant parameters. The specific form of this probability is helpful in characterizing the solution, but it is not restrictive as long as the qualitative properties remain.

The demand of a distressed bank's loans is then determined by the following maximization problem of a potential buyer

$$\max_z \left[(1 + \Psi) \ln \zeta (1 + z) \frac{1 + r}{1 + r_e} \right] (1 + r_e) - (1 + q)z$$

subject to $z \leq f$. The demand for distressed loans is then

$$1 + q_D = \frac{(1 + \Psi)(1 + r)}{1 + z}.$$

The supply of loans is given by the binding liquidity constraint of a distressed inter-

mediary (4), which, given assumptions 1 and 2, can be rewritten as $z(1+q) + (1-f)(1+r) = (1+r)$. Then the supply of distressed loans is

$$1 + q_S = \frac{f(1+r)}{z}.$$

Market clearing implies that $q_D = q_S$. Thus $z^* = \frac{f}{1+\Psi-f}$, subject to the constraint that $z^* \leq f$, which implies,

$$f \leq \Psi. \tag{6}$$

The operation of fire-sale markets puts a bound on the fraction of loans that a bank can hold in order to guarantee enough funds for liquidation in case of distress.

Given these assumptions, a bank simply chooses the fraction f^* of investments in loans and the interest rate r^* to pay to savers, taking as given the securitization technology Ψ and the return r_e. The next proposition summarizes these optimal choices.

Proposition 1 *Banking Optimal Choices.*

The fraction of loans in the portfolio f^* is given by

$$f^* = \min\{1, \Psi\}$$

The payment to savers r^* is given by

$$r^* = r_e - \frac{\widehat{\phi}}{f^*}$$

where f^* and r^* are both increasing in securitization (decreasing in Ψ).

Proof When $r_e > \widehat{\phi}$ the objective is to maximize f subject to the liquidity constraint (4), which in a fire sale market is simply given by constraint (6). Given f^*, the promise to savers, r^*, is determined by the zero profit condition (5). It is trivial that both f^* and r^* are increasing in securitization (decreasing in Ψ).   Q.E.D.

Intuitively, when it is easy to trade assets (a liquid interbank market), there are fewer losses in case of liquidation and distress. The higher is the fire-sale price, the higher is the fraction of loans that a bank can hold and still successfully face a bank run (a higher Ψ allows for a higher f^*). As banks can hold more productive assets in their

15

portfolio and still avoid a run on path, zero profit conditions imply a better return for depositors (a higher f^* implies a higher r^*). Combining equilibrium values of f^* and r^* we can define a risk-adjusted interest spread as

$$\phi \equiv r_e - r^* = \max\left\{\widehat{\phi}, \frac{\widehat{\phi}}{\Psi}\right\}. \tag{7}$$

The risk-adjusted interest spread has two main components: i) the physical cost of production, represented by the *value-added component*, $\widehat{\phi}$ and ii) the *liquidity-premium component*. This last component depends on the securitization technology. It increases as Ψ decreases (securitization becomes worse) and it is zero when $\Psi \geq 1$.

Notice that in this model the liquidity constraint always holds but never binds, which implies that there is never a run in equilibrium and fire sales restrict outcome off-equilibrium. The absence of runs on the equilibrium path is an artifice from the absence of exogenous shocks that force the constraint to bind. This could be easily accommodated, but our intention is to characterize steady states and not fluctuations.

Traditional and Shadow Banks: We assume there are two technologies available in the economy that differ in how loans are packaged, pooled, and tranched to be exchanged in the interbank market. Traditional banks operate with Ψ_{TB} and shadow banks with $\Psi_{SB} > \Psi_{TB}$. First, empirically shadow banks (mutual funds) trade securities much more than traditional banks (commercial banks). Second, traditional banks face larger regulatory constraints, which put exogenous additional constraints on f. This setting justifies shadow banks investing a larger fraction of their portfolio in productive loans, facing less liquidation costs and offering larger return to depositors.

2.1.5 Aggregates and Definition of (Stationary) Equilibrium

Since η, γ, τ and g are all constant in our setting, in what follows we will focus on a balance growth path equilibrium. Along the balanced growth path, all aggregate variables, except L, grow at the rate $\hat{\gamma} = (1 + \gamma)(1 + \eta) - 1$ and all per capita variables grow at the rate γ. For instance, $K_{t+1} = (1+\hat{\gamma})K_t$, while investment is $X_t = (\delta_k + \hat{\gamma})K_t$; therefore, from now on, we omit the time subscript. Even though we will present the main results comparing stationary equilibria, in Section 4.3 we compute the transitions between equilibria.

In a balanced growth path we only need to analyze the problem of an individual born at $t = 0$, as the problem of any other individual born at any other calendar period t is simply $c_{t,j} = (1 + \gamma)^t c_{0,j}$. Thus, we solve for the life pattern of consumption of individuals born at $t = 0$ (that is, $c_{0,j}$) and apply it to all agents born at $t > 0$. Then, in the balance growth path, we simply denote the life pattern of consumption as c_j.

First, we specify aggregates along the balanced growth path. As the only source of heterogeneity in the model comes from α, let $\mathcal{A}^i$ be the stationary set of agents α choosing strategy i, $\mu_i(\alpha) = m(\alpha)$ if $\alpha \in \mathcal{A}^i$ and define $\mu_i = \int_{\alpha \in \mathcal{A}^i} m(\alpha) d\alpha$. In every period t, a density $(1 + \eta)^t m(\alpha)$ of agents are born and their survival probabilities are exogenous; then the density of agents of age j and type α choosing strategy i is

$$\mu_j^i(\alpha) = \begin{cases} \frac{\mu_i(\alpha)}{(1+\eta)^{j-t}} & \text{if } j \leq T \\ \frac{(1-\delta)^{j-T-1}\mu_i(\alpha)}{(1+\eta)^{j-t}} & \text{if } j > T. \end{cases}$$

We use these measures to obtain aggregates for each agent type i, as functions of two endogenous state variables: the marginal productivity of capital, r_e, and the bequest obtained by individuals, $\bar{b}$.

$$\mathbb{C}(r_e, \bar{b}) = \sum_{i=S,B} \sum_{j=1}^{\infty} \int c_j^i(r, \bar{b}; \alpha) \mu_j^i(\alpha) d\alpha$$

$$\mathbb{W}^B(r_e, \bar{b}) = \sum_{j=1}^{\infty} \int w_j^B(r, \bar{b}; \alpha) \mu_j^B(\alpha) d\alpha$$

$$\mathbb{W}^C(r_e, \bar{b}) = \sum_{j=1}^{\infty} \int w_j^C(r, \bar{b}; \alpha) \mu_j^C(\alpha) d\alpha$$

$$\mathbb{B}(r_e, \bar{b}) = \sum_{i=C,B} \sum_{j=T+1}^{\infty} \delta \int b_j(r, \bar{b}; \alpha) \mu_{j-1}^i(\alpha) d\alpha$$

$$L_t = \sum_{j=0}^{T-1} (1 + \eta)^{t-j}$$

where $\mathbb{C}$ is aggregate consumption along the balanced growth path; w^B and w^C are the individual net worths for agents following strategies B (banks) and C (capital markets), respectively; $\mathbb{W}^B$ and $\mathbb{W}^C$ are the corresponding aggregates; $\mathbb{B}$ is the aggregate bequest; and L_t is total labor supply available in calendar period t.

Definition 1 *Stationary Equilibrium.*

Given fiscal policies $\{g, ss_i, D^G\}$, a stationary equilibrium is characterized by saving choices $\{\{TB, SB\}, C\}$, individual allocations $\{\underline{c}(\alpha), \underline{w}(\alpha), \underline{b}(\alpha)\}_{\forall \alpha \geq 0}$, prices $\{y, r_e, r_L, r\}$, and aggregate allocations $\{Y, X, K, \mathbb{B}, \mathbb{C}\}$, such that

1. *Given prices $\{y, r_e, r_L, r\}$ and fiscal policies $\{g, ss_i, D^G\}$, the individual allocations $\{\underline{c}(\alpha), \underline{w}(\alpha), \underline{b}(\alpha)\}$ solve the consumer-saver problem for all $\alpha > 0$: households choose their retirement plan and consumption path to maximize utility.*

2. *Banks choose rates to pay and their portfolio allocation to maximize profits.*

3. *Factor prices are equal to marginal productivities.*

4. *The government chooses τ to balance the budget.*

5. *Markets clear:*

   - *Feasibility:* $\qquad\qquad Y = gY + \mathbb{C}(r_e, \bar{b}) + X + \phi[\frac{\mathbb{W}^B(r_e, \bar{b})}{1+r} - D^G].$
   - *Assets market:* $\qquad \frac{\mathbb{W}^B(r_e, \bar{b})}{1+r} + \frac{\mathbb{W}^C(r_e, \bar{b})}{1+r^e} = D^G + K.$
   - *Bequest=inheritance:* $\bar{b} = (1 + \gamma)^{T_I} \mathbb{B}(r_e, \bar{b}).$

2.2 Equilibrium Characterization

We solve the equilibrium backwards. First, we characterize the optimal consumption path of an α-agent *conditional* on saving in capital markets or through banks. Then we solve for its optimal saving strategy.

First, consider the strategy of saving in banks. The following analysis holds regardless of whether the agent chooses to use traditional or shadow banking, which will be determined later from comparing the higher rate r and the higher cost κ of shadow banking. Since the age of death is the only source of uncertainty, and banks can provide insurance against living too long by pooling resources from a continuum of depositors, the optimal contract agents would sign with banks is an *annuity contract*, which guarantees a constant path of consumption after retirement. In this sense we denote saving for retirement in banks as obtaining *safe assets*.

Then, any agent saving in banks maximize the utility in equation (1) subject to equation (2), knowing that consumption after retirement is constant. In the appendix we show that the solution is characterized by:

$$c_j^B = \bar{c}^B \beta^j (1+r)^j v_0^B,$$

$$b_j^B = \alpha \bar{c}^B \beta^j (1+r)^j v_0^B.$$

(8)

for some constant $\bar{c}^B > 0$. Notice that b can be considered as another consumption good, so that intra-temporal optimality imposes $b = \alpha c$. If $\beta(1+r) = 1$ a depositor would consume a constant amount throughout its lifetime and would leave exactly the same bequest, *independently of how long the household lives.*

This consumption plan implies the following pattern of the net worth evolution,

$$
\begin{aligned}
w_0^B &= 0 \\
w_j^B &= (w_{j-1}^B - c_{j-1}^B + (1-\tau)y_j)(1+r), & 1 \le j \le T, \; j \ne T_I \\
w_j^B &= (w_{j-1}^B - c_{j-1}^B + (1-\tau)y_j)(1+r) + \bar{b}, & j = T_I \\
w_j^B &= \sum_{t=0}^{\infty} \frac{(1-\delta)^{t-1}}{(1+r)^t} \left[(1-\delta)c_{j+t} + \delta\alpha b_{j+t} - ss_B y_T \right], & j > T
\end{aligned}
$$

(9)

Intuitively, agents are born with zero wealth, and as they work, they deposit in banks (at a return r) any non-consumed income. At age T_I each household receives an inheritance, which is mostly saved; thus the net worth jumps at this age. After retirement, banks pay according to the contract specified and the net worth for the household is the present value of the contract.

Now we consider the strategy of saving for retirement in capital markets. In this case households must plan how much to save for retirement and how to spend those savings after retirement. This can be considered as two separate problems. We solve it backwards, solving first the problem after retirement.

Since all bequests are accidental $b_j = w_j$ for all $j \ge T$, the problem after retirement solves

$$V(w) = \max\{\log c + (1-\delta)\beta V(w') + \delta\beta\alpha \log w'\}$$

subject to

$$c + \frac{w'}{(1+r_e)} \le w$$

19

where r_e is the risk-adjusted return on equity, the return or this strategy.

Given the assumed functional forms for consumption and bequests, it is straightforward to verify that the value function is logarithmic in w. That is,

$$V(w) = \bar{\nu}_1(\alpha) + \bar{\nu}_2(\alpha) \log w$$

$$\text{with } \bar{\nu}_2(\alpha) = \frac{1 + \alpha\beta\delta}{1 - (1 - \delta)\beta}.$$

The optimal consumption plan and the implicit optimal bequest plan are then

$$c = w/\bar{\nu}_2(\alpha)$$

$$w' = (1 + r_e)(w - c + ss_S y_T). \tag{10}$$

Given this solution after retirement, the optimal plan at entry in the labor force solves

$$\max \sum_{j=0}^{T-1} \beta^j log c_j + \beta^T V(w_T)$$

subject to

$$\sum_{j=0}^{T-1} \frac{c_j}{(1 + r_e)^j} + \frac{w_T}{(1 + r_e)^T} \leq v_0^C$$

with v_0^S given by equation (2). The solution is

$$c_j^C = \bar{c}^C \beta^j (1 + r_e)^j v_0^C, \qquad\qquad j < T \tag{11}$$

$$w_T^C = [1 - \sum_{j=0}^{T-1} \bar{c}^C \beta^j](1 + r_e)^T v_0^C.$$

During working age, the net worth of agents that follow strategy S evolves as

$$w_0^C = 0 \tag{12}$$

$$w_j^C = (w_{j-1}^C - c_{j-1}^C + (1 - \tau)y_j)(1 + r_e), \qquad 1 \leq j \leq T, \ j \neq T_I$$

$$w_j^C = (w_{j-1}^C - c_{j-1}^C + (1 - \tau)y_j)(1 + r_e) + \bar{b}, \qquad j = T_I$$

Two features of this economy are apparent when we compare equations (11) and (10) with equation (8). First, since $r_e > r$, the consumption of agents who save in

20

capital markets grows faster that the consumption of those who save in banks. After retirement, however, the former experience a faster decline in consumption than the latter. In fact, the consumption of agents that save in capital markets converges to zero as the agent lives long enough. This patterns are summarized in Figure 2. The difference in the return of these two strategies also has the same implications for the evolution of net worth across agents with different saving strategies.

Figure 2: Lifetime Pattern of Consumption Under Strategies B and C
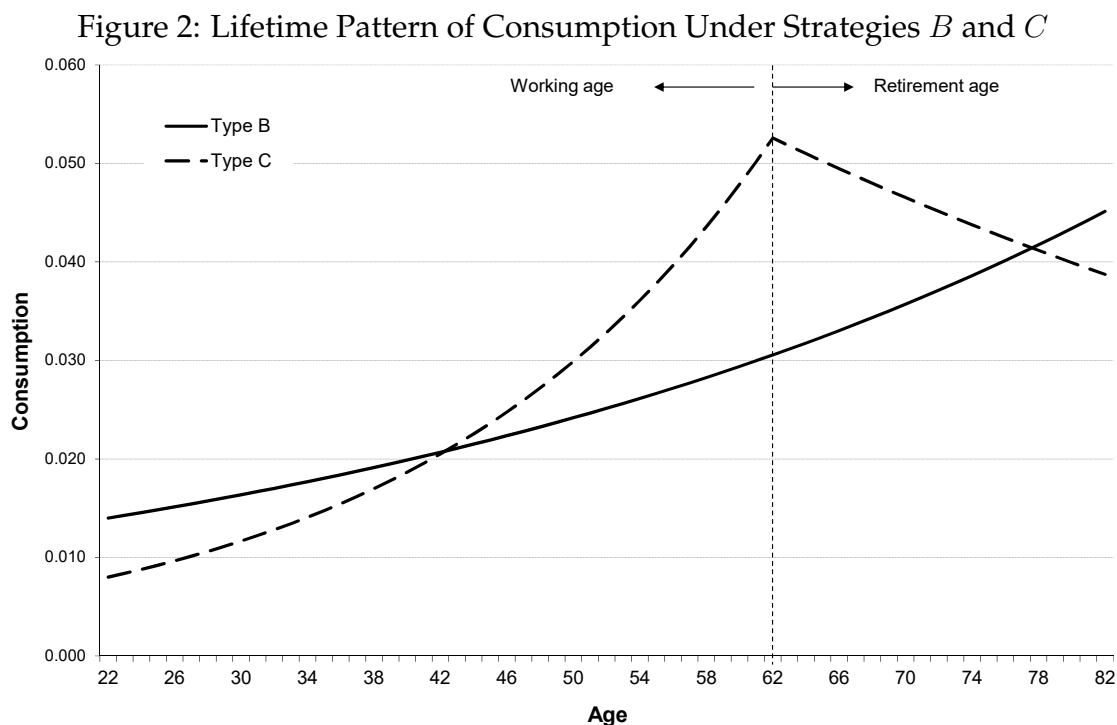
Now, based on these different consumption paths, we characterize the saving strategies of agents with different bequest motives when entering the labor force. First, conditional on depositing in the bank, the agent must choose a traditional or a shadow bank. The trade-off between these two alternatives is that the return from saving in shadow banks is higher, but represents a utility cost κ of searching, understanding the contract, and potentially facing an aggregate crisis, incurred at the time of signing the contract. The next proposition shows that, conditional on depositing in a bank, the agent chooses shadow banking when expects to live long enough, enjoying the additional return longer at the same cost κ. This is true when the agents' bequest motive is not so large. As we show next, these are the agents selecting into banking.

21

Proposition 2 *Choice between traditional and shadow banking.*

For agents with relatively low bequest motives ($\alpha < \frac{1}{1-\beta}$), there exists a unique $\delta^*(\alpha, \kappa) > 0$ such that, when $\delta \geq \delta^*(\alpha, \kappa)$, households that follow strategy B sign the annuity contract with traditional banks, and when $\delta < \delta^*(\alpha, \kappa)$, they sign the annuity contract with shadow banks. Furthermore, $\delta^*(\alpha, \kappa)$ is increasing in α and decreasing in κ.

For this result, it is important that κ is constant and independent of δ. If κ were solely capturing search and attention costs of securitization, this assumption would arise naturally. If in addition κ captures the probability that securitization collapses, it could also depend on fundamental parameters. To address this issue in Appendix B we consider an alternative environment where instead of the fixing κ, agents face a constant annual probability p of loosing wealth equivalent to $(1 - \zeta)$ units of consumption. Then, the microfounded equivalent of κ would be,

$$\kappa(\delta) = -\beta^T \frac{p \log(\zeta)}{1 - \beta(1 - \delta)} > 0.$$

This representation of the cost κ is increasing both in the probability and the losses of a crisis (note that because $\zeta < 1$, then $\log(\zeta) < 0$). The cost is also increasing on life expectancy, but as we show in Appendix B, the benefit of higher rates increases at a faster rate than this cost. Thus, under some additional conditions (also satisfied in our quantitative exercise), we are able to prove an analogous result to Proposition 2.

Once determined what is the bank the agent would use if saving in a bank, agents choose between saving in banks or in capital markets. Saving in banks has the benefit of fully insuring against the risk of living long, but it has the cost of low deposit return. Conversely, saving in capital markets has the benefit of high asset returns, but at the cost of not providing insurance against living too long. In particular, agents saving in capital markets may leave large amounts of accidental bequests. Of course, the stronger is the household's bequest motive the lower the implicit cost of accidental bequests. This intuition is confirmed by the next proposition.

Proposition 3 *Choice between banks and capital markets.*

There are $\bar{\phi} > \underline{\phi} > 0$ such that for all $\widehat{\phi} \in [\underline{\phi}, \bar{\phi}]$, there exists a unique $\alpha^*(\delta) > 0$ such that all agents with $\alpha < \alpha^*(\delta)$ follow strategy B and all agents with $\alpha \geq \alpha^*(\delta)$ follow strategy S.

Note that in this economy all agents have access to banks. That is, to safe assets that deliver the same consumption after retirement, regardless of when the agent dies. Individuals with high bequest motives, however, optimally choose not to use them.[15]

From Proposition 3, the fraction of the population using banks and capital markets depends on the distribution of bequest motives, α. Similarly, from Proposition 2, the fraction of the population using traditional and shadow banks depends both on the distribution of bequest motives, α, and of shadow banking costs, κ. In what follows we make the following assumption about these distributions.

Assumption 3 *We assume that the distribution of bequest motives is concentrated in two points: $\alpha = 0$ with probability μ and $\alpha = \widehat{\alpha} > 0$ with probability $(1 - \mu)$. We also assume a single and fixed cost of shadow banking κ for all agents.*

Since there are only two saving alternatives, banks and capital markets, the first part of the assumption is qualitatively without loss of generality. Agents with $\alpha = 0$ will save in banks (as $r_e \rightarrow r$), hence we need to guarantee that $\widehat{\alpha}$ is high enough to also have agents saving in capital markets. This assumption immediately implies from Proposition 3 that the size of the banking industry is pinned down by μ, which we assume exogenous. Endogenizing the size of the banking sector is in part the motivation of Scharfstein (2018), but beyond our scope, which instead focuses on the change in composition *within* the banking system.

The combination of the two parts of the assumption has, however, implications for the composition of the banking industry. Since all agents saving in banks (those with $\alpha = 0$) face the same κ, the threshold $\delta^*(\alpha, \kappa)$ from Proposition 2 is identical for all of them, and then all choose to switch to shadow banks simultaneously. This simplification is motivated by the difficulty to measure intermediation costs of traditional and shadow banks separately, which forces us to target *average intermediation costs* when performing the calibration. Since it is irrelevant in the aggregate whether the observed reduction in average intermediation costs comes from a wide adoption of shadow banking with slight less intermediation costs or by a moderate adoption of shadow banks with much lower intermediation costs, we assume the first case, with all agents adopting shadow banking at the same time. There is always, however, a distribution of κ that can perfectly match the evolution of shadow bank adoption.

---

[15]This mechanism is in line with the recent finding by Lockwood (2012 and 2015), who argues that a high bequest motive could be an explanation for the "annuity puzzle".

3 Measuring Shadow Banking and Intermediation Costs

In this section, and in preparation to evaluate the model quantitatively, we document the evolution of average intermediation costs since 1980 and discuss the role of shadow banking in interpreting such evolution.

As there is no readily available measure of ϕ in the aggregate, as a proxy for intermediation costs we use spreads between total interests received and total interests paid in the whole financial sector, from NIPA tables. We have to make some adjustments, however. First, productive investment opportunities are risky and some loans are not recovered by the bank. Hence, in the data, we subtract from interests received the "bad debt expenses."[16] Second, there are many services provided by banks that are not priced in, such as safety, accessibility to ATMs, financial advising, insurance, etc. Thus, we add to the interests paid by the financial sector the "services furnished without payment," which assigns a monetary value to these services.

To be more precise we want to measure $\phi = r_e - r$, where r_e has to be corrected for defaulting debt and r for non-priced services. We can decompose ϕ into measurable components as

$$\phi = r_e - (\overbrace{r_L + r_s}^{r}) = \frac{\overbrace{f r^e + (1 - f) r_L}^{r_T} - r_L}{f} - r_s,$$

where r_L is the interest paid for deposits (same as bond returns), r_s is the return for other services not priced by banks, f the fraction of portfolios in productive loans and $r_e = (1 - s_b)(1 + \widehat{r}_e) - 1$, with $\widehat{r}_e$ being the rate charged for loans and s_b the fraction that defaults. These components have counterparts in NIPA tables:

1. r_T=(Total interest received - bad debt expenses)/hh's debt.

   This expression measures the average return on assets for all concepts that banks receive. We use Table 7.11, Line 28 of the NIPA tables, which provides the total interest received by private financial intermediaries and subtract Table 7.1.6 Line 12 of the NIPA table that provides "bad debt expenses" declared by corporate business.[17] To express

---

[16]If the law of large numbers holds for financial institutions, the average loss per unit should be equal to the average. From this point of view, the adjusted interest received could be considered as the equivalent risk free return on loans.

[17]As not all corporate business are financial intermediaries, we follow Mehra, Piguillem, and Prescott (2011) and assign half of it to the financial sector. We also perform alternative calculations assigning 25%, 75% and 100% to the financial sector without any qualitative change, just a change in levels.

these values as returns, we divide them by all household liabilities (*hh's debt*) from Table D.3 of the Flow of Funds.

2. r_L=(Total interest paid)/hh's debt.

   This expression measures the average return on deposits that depositors and savers receive. Table 7.11, Line 4 of NIPA provides information about the total interest paid on deposits by the financial sector, which we divide by *hh's debt*.

3. r_s=(Services furnished without payment)/hh's debt.

   This expression measures the average return on services provided by financial intermediaries that are not explicitly charged to depositors and savers. We obtain this figure from Table 2.4.5, Line 88 of the Flow of Funds, which we also divide by *hh's debt*.

4. f=Fraction of portfolio of financial intermediaries allocated to productive investments

   This is perhaps the most difficult figure to measure, but also central to our analysis. We denote by s the fraction of intermediaries not chartered as depository institutions, call them *shadow banking institutions* (hedge funds, SIVs, investment banks, money market funds, etc.) and assume they allocate all of their portfolio to productive assets. The remaining fraction corresponds to *traditional banks* that only allocate a fraction $\widehat{f}$ of their assets to productive assets (either they suffer the threat of runs or they are regulation constrained). The fraction of productive investments in the financial sector is then,

   $$f = s + (1 - s)\widehat{f}.$$

   Measuring s is challenging because part of traditional banks' investments is also channeled by *shadow banking activities* (securitization, sponsoring of special purpose vehicles, participation in repo markets, etc.). To avoid double counting and taking a stand on what should be classified or not, instead of measuring shadow banking directly we measure it as a residual from traditional baking activities. First, we compute $(1 - s)$ by the fraction of consumer credit and mortgages to households that is channeled through traditional banks (from Table 110 we divide consumer credit from Line 14 plus mortgages from Line 15 by total consumer credit and mortgages obtained by all households from Table D3). Then, we compute $\widehat{f}$ by the fraction of loans in the portfolio of traditional banks (from Table 110 we divide all the loans from traditional banks from Lines 12, 14 and 15 by all their deposits, checkable and savings, from Lines 23 and 24).

Combining these components, Figure 3 shows the spreads since the seventies. In short, right before 1980 spreads were stable at around 4%, there was an increase in

Figure 3: Risk-Adjusted Spread, ϕ
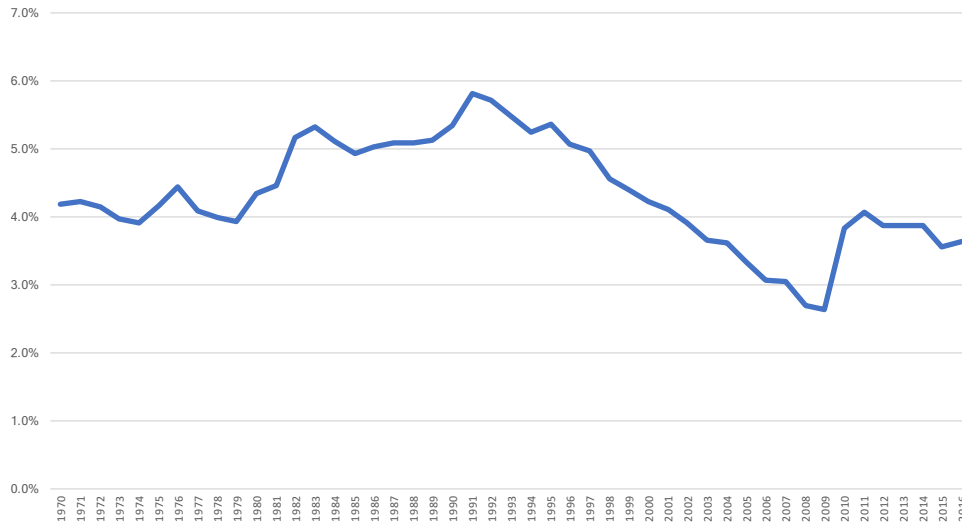
the 80s and 90s, and a large decline that reached 3% before the great recession, to jump again in recent years to pre-1980s levels.

Figures 4 and 5 show the decomposition of the spread. According to our measures, shadow banking institutions (s) increased from 5% in the seventies to more than 50% in recent years, while shadow banking activities of traditional institutions (captured in part by $\widehat{f}$) also increased from 80% in the seventies to almost 100% before the crisis, and then collapsed to 70% right after the recent U.S. financial crisis.

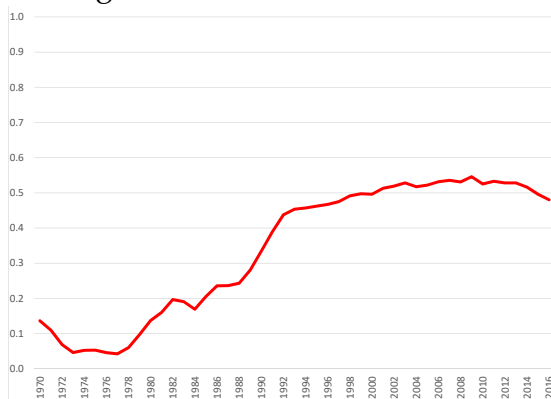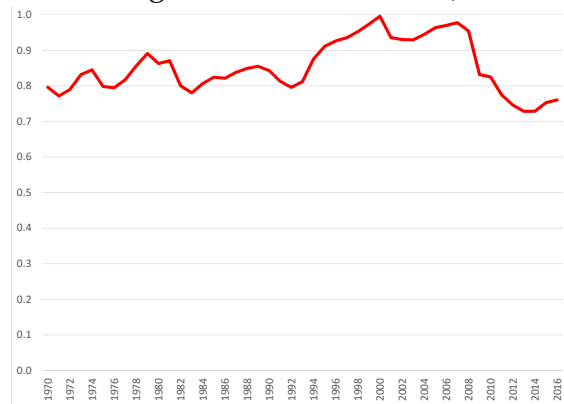

Figure 4: SB Intermediaries, s

Figure 5: SB Activities, $\widehat{f}$

Why did spreads decline in the last decades? Have financial intermediaries improved either their management efficiency or their asset liquidation value? Philippon (2015)

performs a thorough calculation of the changes in efficiency of the financial sector in the U.S. during the last century using data on value added. He shows that the technology in the financial intermediation industry exhibits constant returns to scale and $\widehat{\phi}$ has been constant at roughly 2% for more than 100 years.[18]

This result implies that the liquidity premium accounts for most of the observed variation in the risk-adjusted spread. To see this, we define the liquidity premium as

$$\text{Liquidity premium} = (1 - f)(r_e - r_L)$$

This is the difference between the realized spread and the spread if liquidity were not an issue. Figure 6 shows the evolution of this premium during since 1970, which declined from around 1% to almost 0% by 2007. After the recent financial crisis, the liquidity premium of intermediation increased again to almost 0.5%. The pattern in

Figure 6: Liquidity Premium
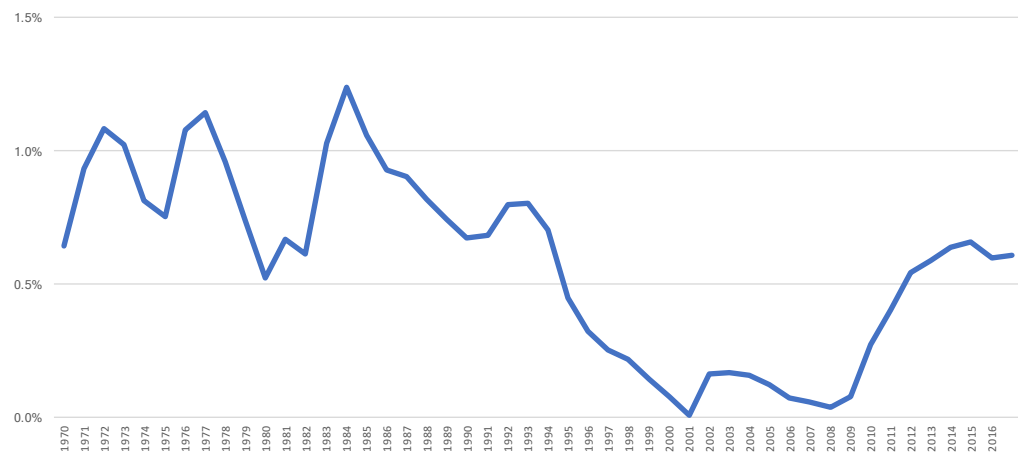
Figure 6 is surprisingly similar to the pattern documented by Del Negro et al. (2017b) in the same timeframe, with a completely different methodology. They also show that 1 percentage point of the fall in the U.S. natural interest rate experienced from the eighties until the great recession is explained by the change in the convenience

---

[18]Philippon (2015) performs two alternative calculations: one assuming that the composition of the types of loans offered by the financial sector was stable during the sample period and another adjusting for changes in the quality of the loans. When computing the per-unit value added, Philippon (2015) explicitly, and correctly, discards the use of intermediation spreads as measures of value added. As we show in equation (7), intermediation spreads are affected by other factors that, even though not reflect physical costs, deeply affect the cost of financial intermediation.

yield. Further, Del Negro et al. (2017a) point to the relevant role played by shadow banks as liquidity providers.

In short, the three decades before the recent crisis was characterized by a large drop in the financial intermediation spread, almost exclusively led by a reduction in the financial sector's "liquidity premium." Shadow banking has had a direct impact by improving the assets' tradability, replacing less profitable government bonds on banks' balance sheets by more productive loans, and then improving investment and output.

4   Quantitative Assessment of the Model

To decompose the direct macroeconomic effects of an increase in life expectancy and its indirect effect through transforming the financial system towards more securitization and lower intermediation costs, we first calibrate the economy to replicate the main aggregates for financial intermediation in 1980. Then, we obtain the model's output for 2007 keeping most of the parametrization as in 1980 and we only change the newly observed life expectancy and intermediation costs. We also analyze what would have happened if the United States had to face the demographic transition while forbidding the use of securitization and shadow banking.

4.1   Calibration for 1980

We calibrate the model to yearly data. There are some parameters that are standard in the literature: i) the discount factor $\beta = 0.9975$, ii) capital share $\theta = 0.33$ consistent with a capital income share of output equal to 33%, iii) labor productivity growth, $\gamma = 0.02$ and iv) population growth, $\eta = 0.01$.[19]

When choosing the capital depreciation rate, δ_k, we need to take into account what capital means in our economy. The literature targets a capital-output ratio of 2.7,

---

[19]The calibrated β parameter is larger than standard values in the literature. The reason is that, for tractability, we have fixed the coefficient of relative risk aversion parameter to 1 (log preferences) and then the discount factor must capture how agents assess, when entering the labor force, the risk of death after retirement. Most of the discounting comes from the probability of death, which is absent in most macroeconomic models. Indeed, Giglio, Maggiori, and Stroebel (2015) find that households barely discount payoffs that happen in the very long run. In any event, using preferences with arbitrary relative risk aversion would allows us to decrease the value of β by increasing the coefficient of risk aversion. See also the discussion in Krueger and Kubler (2005).

which is the approximate ratio for the U.S. economy when one focuses only on productive capital. In our economy capital includes many other physical assets that constitute wealth for the household, such as housing and land. Including these assets capital-output ratio is about 3.4, which is generated by a depreciation rate of $\delta_k = 0.0271$. Regarding the agents' lifecycle, we assume that they enter the labor force at age 23, retire at age 63 (this is, $T = 40$) and, based on the Survey of Consumer Finances, receive inheritance at age 52 (this is, $T_I = 29$).

Regarding μ (the fraction of agents with $\alpha = 0$ who choose to save in banks), we discipline it with the fraction of agents directly participating in capital markets. Thus, we measure it with the fraction of financial assets held in corporate and non-corporate equity from the Flow of Funds (the first panel of Figure 1). This figure has been roughly constant in the U.S. at around 28% since the eighties. Accordingly, we calibrate $\mu = 0.72$ such that $1 - \mu = 0.28$.

As for the parameters that determine fiscal policies, according to NIPA government spending is 20% of GDP (this is, $g = 0.20$) and in 1980 government debt (federal, state and local) was around 40% of GDP. Since we model a closed economy and around 20% of the government debt was held by foreign investors, we set $D^G/Y = 0.33$.

Finally, two parameters remain to be calibrated: i) the bequest motive, $\widehat{\alpha}$, and ii) the fraction of the last wage that the government transfers as Social Security after retirement, ss_i. Since there is no direct measure of these parameters, we normalize $ss_C = 0$ (no social security for investors in capital markets) and choose ss_B and $\widehat{\alpha}$ to replicate two moments in the data: i) government debt to GDP ratio of 0.33 in 1980 and ii) household debt to GDP ratio of 1 in 1980. This implies $\widehat{\alpha} = 4.64$ and $ss_B = 0.55$. To assess the validity of these parameters notice that: i) $\widehat{\alpha}$ of around 4.6 generates in the model a level of savings consistent with the findings from De Nardi, French, and Jones (2015) and ii) ss_B of 55% of the last wage implies a ratio of Social Security of 34% of the average wage, which is consistent with information from the Social Security Administration.[20]

The two counterfactual parameters that we will modify between 1980 and 2007 are: i) the survival probability after retirement, δ, which captures life expectancy and ii) the

---

[20]Monthly average payments per retired beneficiary were around $1,250 per month in 2015. Given an average annual wage of $57,000 in 2014, this implies a ratio of 27%, which is lower than the ratio generated by the model. Including Medicare and Medicaid, however, would raise this ratio closer to the one implied by the model.

spread between borrowing and lending, ϕ, which captures the role of securitization and shadow banking to reduce intermediation costs. We start by calibrating $\delta = 0.072$ for 1980, which implies a life expectancy of 13.9 years after retirement. In the counterfactual we decrease this value to $\delta = 0.052$, which implies a life expectancy of 19.23 years after retirement, the observed value in 2007.[21] Based on Section 3, we calibrate $\phi = 0.04$ for 1980 and in the counterfactual we decrease its value to $\phi = 0.03$, which is the observed value in 2007.

We can assess the model's performance for moments that we have not targeted. First, the model generates a ratio of private consumption to GDP of 0.56 in 1980, close to the observed ratio of 0.62. Second, the model is not as successful on capturing the amount of inheritances, as it generates 4.9% of GDP, while most empirical studies estimate this figure to be around 2.7%. Those empirical estimates, however, abstract from inter-vivos transfers whose present value could be larger than the inheritance.

Remark on savings rates: Given the long standing debate about the declining "savings rate" in the U.S. (see Summers and Carroll (1987)), it may appear puzzling that our theory relies on larger savings to explain the increased in financial intermediation. This apparent contradiction can be settled, however, by properly defining "savings". While such literature documents a sharp decline in the *flow of savings* out of disposable income, in most theories, including ours, savings are defined at a given time by the *stock of financial assets*. There are two main reasons why these two measures could generate contradictory results. First, in short windows larger savings needs (in terms of stocks) are perfectly compatible with lower savings flows whenever an economy is in a transition or affected by shocks. Second, the stocks of financial assets are greatly affected by the prices of those assets. Thus, as long as individuals do not consumed their capital gains, it is possible to increase the stock of savings with declining (or even negative) flows of savings. The relevance of this last channel has been stressed by Fagereng et al. (2019) to explain the behavior of savings rates in Sweden. The fact that in the U.S. economy both the capital-output ratio and the wealth to GDP-ratio have sharply increased since 1980 (see next section) confirms our theory once we define savings based on the stock of financial assets.

---

[21]These numbers can be computed using the actuarial tables of the Social Security Administration given that workers retire at an average age of 63. We use slightly conservative estimates, since life expectancy is still in a downward path.

4.2 Decomposing the Role of Life Expectancy and Shadow Banking

We now perform a counterfactual exercise, decomposing the effects of the change of life expectancy and the change in intermediation costs on asset accumulation, output and welfare, from 1980 to 2007, before the crisis.

What parameters do we use for the counterfactual on 2007? Most parameters have not changed, but some have. First, the population growth rate increased to 1.4% in 1992, and then fell to 0.7% in 2011. In our counterfactual we set $\eta = 0.007$ for 2007. Second, we maintain a government debt of 33% as a ratio of GDP in 2007. Even though the ratio increased to 62%, around 45% of U.S. federal debt was held by non-U.S. residents.[22] Based on these figures we argue that the increase in the provision of government bonds during this period was not relevant domestically. Third, as in the data, we maintain the replacement ratio (that is, the proportion of wages obtained from the government after retirement) and allow labor taxes to adjust in order to satisfy the government budget constraint. Later, we show the same simulations, but we keep the labor tax constant and allow the government debt to change. This last exercise is helpful to understand the underlying mechanisms affecting our results.

In Table 1, the first column shows the calibration results for 1980. The last column introduces the counterfactual when life expectancy increases (captured by a reduction in δ from 0.072 to 0.052) and agents move bank's savings from traditional to shadow banks. Because of Proposition 1, there are two levels of utility costs to sign a contract with securitization, $0 < \underline{\kappa} < \bar{\kappa}$ such that if $\kappa \in [\underline{\kappa}, \bar{\kappa}]$, it is optimal for agents to choose traditional banks when $\delta = 0.072$ and shadow banks when $\delta = 0.052$. Due to the move from traditional to shadow banks, the intermediation spread falls from $\phi = 0.04$ to $\phi = 0.03$, as we observe in the data and the model in Section 3.

Comparing the first and last columns, in which we allow for both an increase in life expectancy and a reduction in spreads, the model generates a large increase in the output steady-state level (of around 7%), an increase in the capital to output ratio (from 3.4 to 3.9) and a large increase in households' total financial assets (from 1.33 to 1.94 of GDP). While the data counterparts of the first two figures are difficult to observe, we use Table L100 of the Flow of Funds to measure the increase of households'

---

[22]See http://www.treasury.gov/resource-center/data-chart-center/tic/Pages/ticsec2.aspx. See also Bertaut et al. (2012) for a detailed discussion about the international savings glut in the U.S. economy.

31

financial assets. Subtracting from the total domestic non-financial assets (Line 1, Table L100) the corporate equity (Line 16, Table L100) and the equity on non-corporate businesses (Line 23, Table L100), we obtain a proxy for the net worth of households that use intermediation, which grew from 1.36 of GDP to 2.33 of GDP, very close to the model's prediction. Finally, the model's prediction of the change in the new amount intermediated, measured by the household debt to GDP ratio, accounts for more than 90% of the observed change (the model generates 1.62 as opposed to the 1.66 measured in the data).

Table 1: Counterfactual to 2007 (Fixed D^G)

Economy	1980 Benchmark	Larger δ $\kappa > \bar{\kappa}$	Same δ $\kappa < \underline{\kappa}$	δ & ϕ change $\kappa \in [\underline{\kappa}, \bar{\kappa}]$
Interm. Cost (ϕ)	4%	4%	3%	3%
Survival prob. (δ)	0.072	0.052	0.072	0.052
Interest Rates				
Borrowing Rate (r)	0.030	0.023	0.034	0.028
Lending Rate (r_e)	0.070	0.063	0.064	0.058
National Accounts				
Output	1.00	1.035	1.031	1.070
Capital to output ratio	3.40	3.65	3.62	3.90
Net Worth				
Total	3.73	3.98	3.95	4.23
Equity (Plan C)	2.40	2.68	2.08	2.28
Bank Debt (Plan B)	1.33	1.30	1.86	1.94
Data (FF: Table L100)	*1.36*			2.33
Bequest/Y	0.049	0.049	0.040	0.039
Government Debt/Y	0.33	0.33	0.33	0.33
Household Debt/Y	1.00	0.96	1.53	1.62
Data (FF: Table D3)	*1.00*			*1.66*
Change in welfare at birth	-	-	0.3%	0.4%
Plan C	-	-	-4.3%	-4.8%
Plan B	-	-	2.5%	2.8%

Now we can decompose the effects of the increase in life expectancy and the decline in intermediation costs by suppressing one at a time. The second column of Table 1 shows the counterfactual *without shadow banks*. We compute the model with life expectancy increasing in the same magnitude as observed in the data, but assuming that $\kappa > \bar{\kappa}$, so that the migration toward shadow banking does not happen and, without securitization, spreads remain at 1980 levels. In this case the increase in the capital to output ratio and steady state output would have been around 50% of the total increase with the presence of shadow banking (the capital to output ratio would have increased from 3.4 to 3.65 instead of to 3.9, while output would have increased from 1 to 1.035 instead of to 1.07). Also, absent shadow banking we would have not observed any change in the net worth held by agents saving through banks, in terms of GDP (roughly constant at 1.3), nor in household debt over GDP (roughly constant at 1). Finally, the increase in retirement needs without an improvement in intermediation costs would have increased the demand for savings without an increase in supply, generating a reduction in savings return (r declines from 3% to 2.3%). Still, since there are more funds channeled to investment opportunities, the equity return declines (r_e declines from 7% to 6.3%).

Finally, the third column of Table 1 is a thought experiment *without an increase in life expectancy*, where we assume that κ falls below the lower bound $\underline{\kappa}$, still inducing a movement towards shadow banking. With shadow banking but no extra needs for retirement, the increase in capital to output ratio and steady state output would have been between 40% and 50% of the total increase with higher retirement needs (the capital to output ratio would have increased from 3.4 to 3.62 instead of to 3.9, while output would have increased from 1 to 1.031 instead of to 1.07). That is, the arrival of shadow banking without an increase in the demand of savings would have generated a permanent increase in GDP of almost 3% instead of 7%. We would have observed, however, a large increase in the net worth held by agents in terms of GDP (from 1.33 to 1.86 instead of to 1.94) and household debt over GDP (from 1 to 1.53 instead of to 1.62), almost accounting for the full observed change. Finally, the increase in the supply of savings without an increase in the demand for savings induces an increase in savings returns (r increases from 3% to 3.4%). Still, since more funds are channeled to investments, the equity return still declines (r_e declines from 7% to 6.4%).

Partial equilibrium intuition of the decomposition results: To build intuition about the forces behind the previous decomposition, we show in Figure 7 the partial equi-

librium effects of changes in both life expectancy and intermediation costs on interest rates, capital and credit.

In the left panel we depict the equilibrium in *capital markets*. The decreasing solid line shows the supply of capital, which is the level of K that technologically satisfies: $r_e = f'(K) - \delta_k$. For this reason, it will not be affected by changes in either demographics or financial technology. The increasing solid curve with dot markers is the demand for capital. This can be decomposed between the *direct demand with own funds* by C-agents (the net worth that agents of type C are willing to accumulate at a given interest rate r_e, given by $\mathbb{W}^C(r_e)/(1+r_e)$,) depicted as the increasing solid curve without markers. As the return on capital r_e increases, agents want to increase their savings, demanding more capital. The second component is the *indirect demand with borrowed funds*, which is the amount of funds that banks channel to C-agents to buy capital. This second component, however, is determined by the operation of banks, which we denote next as the credit market.

In the right panel of Figure 7 we depict the equilibrium in *credit markets*. The credit supply, depicted by the increasing solid blue function, is given by the net worth of B-agents and accumulate at an interest rate r, not held in government bonds (that is, $\mathbb{W}^B(r)/(1+r) - D^G(r)$). The solid decreasing red curve, $\Upsilon(r+\phi) \equiv K(r+\phi) - \mathbb{W}^C(r+\phi)/(1+r+\phi)$, is the demand for credit to buy capital, which is the capital that cannot be bought by C-agents with their own funds. As is clear, both markets are interlinked and cannot be solved separately.

The solid lines in both panels of Figure 7 are computed assuming $\delta = 0.072$, as in the first column of the 1980 benchmark. Equilibria in both markets are represented by $E0$. In credit markets this implies a ratio of debt over GDP of 1 and $r = 0.03$. In capital markets this implies a capital to output ratio of 3.4 and a return on capital of $r_e = 0.07$. These are the results in the first column of Table 1, consistent with $\phi = r_e - r = 0.04$.

What happens in this partial equilibrium analysis when life expectancy increases? This counterfactual is shown with dotted lines. First, we discuss what happens in credit markets. Since B-agents expect to live longer, they accumulate more assets in banks, increasing the supply of credit in the economy. C-agents also expect to live longer and assign more of their own wages to buy stocks, reducing the demand of credit. As a result, the *new partial equilibrium* is at point B, with approximately the same amount of private debt (around 1), but with a much lower credit rate ($r = 0.022$).

34

Figure 7: Partial effects
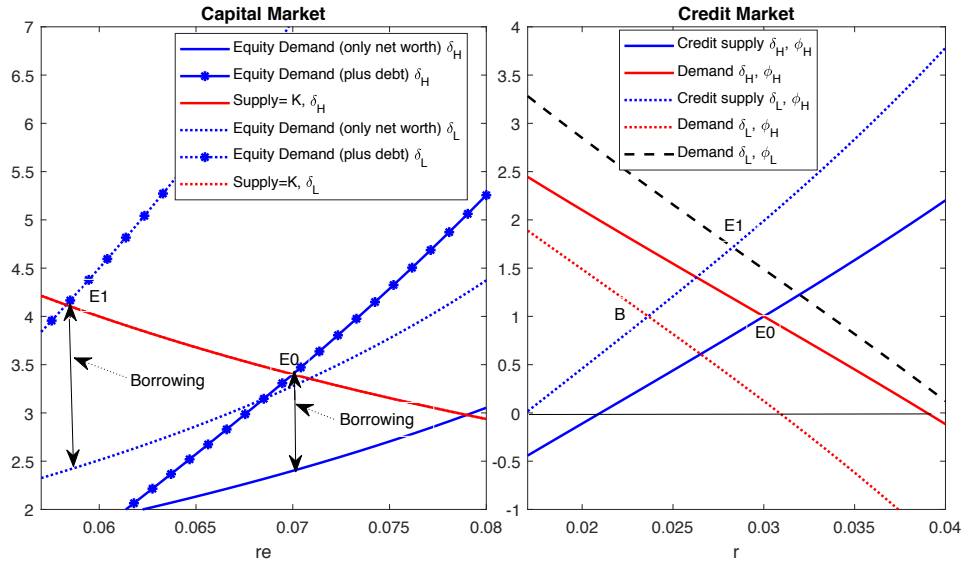
What happens in capital markets? While the supply of capital is not affected (it is purely a technological function), the demand increases because C-agents save more.

Intuitively, the higher demand of savings for retirement by all agents in the economy counteract each other, generating an increase in both the demand and supply for credit, reducing the return on safe assets and increasing capital in the economy, but not changing the total amount of credit.

The arrival of shadow banking tends to have a first-order effect on increasing credit in the economy. A fall in ϕ does not directly affect capital markets, but it does affect the functioning of credit markets. A fall in intermediation costs reduces the cost of credit for C-agents, increasing their demand of credit. The new "partial" equilibrium is at point $E1$, with a higher intermediate interest rate r, more credit and an even higher level of capital, all of which could not be generated by just changing δ.

We emphasize the partial nature of this intuitive analysis, as changes in these two markets will feedback to each other via the quantity of capital in the economy and the accumulation of net worth. In particular, the slopes of demand and supply functions depend on general equilibrium forces. These general equilibrium effects are fully accounted for in the table, but the figure is useful to understand the underlying mechanisms and interactions between agents.

35

Remarks on welfare effects: When there are changes to "preferences" (in our case life expectancy) affecting the computation of present values, comparisons across different scenarios are hard to interpret in terms of welfare. Still, we can make comparisons using consumption equivalent changes when fixing δ.

Let $\underline{c} = \{c_t, b_t\}_{t=0}^{\infty}$ be the sequence of consumption and bequest for an agent at birth *before* a change in the economy and $\underline{\widetilde{c}} = \{\tilde{c}_t, \tilde{b}_t\}_{t=0}^{\infty}$ the analogous sequence *after* the change. We define the consumption equivalent parameter λ as the constant proportional change in every period allocation that makes the consumer indifferent between the two scenarios. That is, λ solves $\sum_{t=0}^{\infty}((1 - \delta_t)\beta)^t u((1 + \lambda)c_t, (1 + \lambda)b_t) = \sum_{t=0}^{\infty}((1 - \delta_t)\beta)^t u(\tilde{c}_t, \tilde{b}_t)$. If λ is positive the consumer benefits from the change, while, if it is negative, the consumer is worse off, since preferences are logarithmic. The above equation can be written as $\sum_{t=0}^{\infty}((1 - \delta_t)\beta)^t \log(1 + \lambda) + \sum_{t=0}^{\infty}((1 - \delta_t)\beta)^t u(c_t, b_t) = \sum_{t=0}^{\infty}((1 - \delta_t)\beta)^t u(\tilde{c}_t, \tilde{b}_t)$. Let $U_0(c)$ be the utility at birth, then λ satisfies:

$$\left[\frac{1 - \beta^{T+1}}{1 - \beta} + \frac{\beta^T}{1 - \beta(1 - \delta)} \right] \log(1 + \lambda) = U_0(\underline{\widetilde{c}}) - U_0(\underline{c})$$

$$\lambda = \exp\left[-\frac{1 - \beta^{T+1}}{1 - \beta} - \frac{\beta^T}{1 - \beta(1 - \delta)} \right] \exp\left[U_0(\underline{\widetilde{c}}) - U_0(\underline{c}) \right] - 1$$

Comparing columns 1 and 3 of Table 1, where the economy has the same $\delta = 0.072$ but lower intermediation costs due to shadow banking, we observe a net increase in welfare of 0.3%. This increase, however, is not without redistribution consequences. While B-agents (who represent almost 70% of the agents) experience a consumption equivalent increase of 2.5%, C-agents experience a drastic decrease of 4.3%. Comparing columns 2 and 4, the economy has a higher life expectancy (of $\delta = 0.052$) and the same reduction in intermediation costs. Shadow banking improves welfare by 30% more when agents live longer (from 0.3% to 0.4%), with stronger redistribution consequences.

Remarks on flexible government debt: In the previous simulations we have maintained D^G fixed. In Table 2 we consider alternative scenarios, with changing D^G. The first column just replicates the calibration in Table 1, while the second column replicates the counterfactual for 2007 when allowing both retirement needs and intermediation costs to vary (the last column of Table 1). The third column shows what the equilibrium would have been if life expectancy had increased, the spread had

decreased to 3% and the government were allowed to freely choose the level of debt without changing taxes. In this case, the government would have chosen a similar level of tax-debt combination, which is due to the similar expenses due to the social Security System. As a consequence, in this scenario the main variables would have remained very similar to just fixing debt to the 1980 level.

Table 2: Counterfactual to 2007 (alternative D^G)

Economy	1980 Benchmark	2007 Calibration	Free D^G	All D^G Domestic
Interm. Cost (ϕ)	4%	3%	3%	3%
Survival prob. (δ)	0.072	0.052	0.052	0.052
Interest Rates				
Borrowing Rate (r)	0.030	0.028	0.027	0.029
Lending Rate (r_e)	0.070	0.058	0.057	0.059
National Accounts				
Output	1.00	1.070	1.071	1.060
Capital to output ratio	3.40	3.90	3.91	3.85
Net Worth				
Total	3.73	4.23	4.21	4.47
Equity (Plan C)	2.40	2.28	2.28	2.36
Debt (Plan B)	1.33	1.94	1.93	2.11
Data (FF: Table L100)	*1.36*	*2.33*		
Bequest/Y	0.049	0.039	0.039	0.041
Government Debt/Y	0.33	0.33	0.30	0.62
Household Debt/Y	1.00	1.62	1.63	1.49
Data (FF: Table D3)	*1.00*	*1.66*		

By changing the government debt to GDP ratio, we can also shed light on what would have happened in the U.S. without an international saving glut contemporaneous with the domestic savings glut. Justiniano et al. (2013 and 2015) relate the credit boom experienced before the crisis to the international savings glut, claiming the fall in interest rates and between a fourth and a third of the higher U.S. household debt can be attributed to the influx of foreign funds. The last column assumes that the debt

to GDP ratio moves from 0.33 (as in 1980) to 0.62, the domestic supply of government bonds in 2007 if foreign nations were not holding any U.S. Treasuries.

The direct effect of more government debt is an increase in interest rates, r, by 20 basis points. This result is consistent with half the estimate of the elasticity of U.S. Treasury yields to U.S. government debt provided by Krishnamurthy and Vissing-Jorgensen (2012) (they estimate that doubling U.S. government debt roughly increases Treasury interest rates by 40 basis points).[23] This change in interest rates induces a decline in private credit (household debt to GDP ratio) with respect to the case in which there is no global savings glut, to 1.49GDP instead of 1.62GDP. This result implies that the international demand for U.S. Treasuries would account for around 21% of the generated increased in the credit boom. This number is very close to the interval provided by Justiniano, Primiceri, and Tambalotti (2013) for the contribution of the international savings glut to the credit boom in the 2000s. However, in our setup the channel is different. There is no *direct* supply of foreign funds (lenders) generating incentives that stimulate households borrowing. Instead, the foreign demand for U.S. Treasuries crowds out the domestic demand for safe assets.

In summary, without a foreign savings glut, the U.S. economy would have experienced a smaller increase in capital-output ratio and in output (on the order of 15% lower steady state output), as there would have been a larger supply of safe assets that forced an increase in the return on capital and less investment.
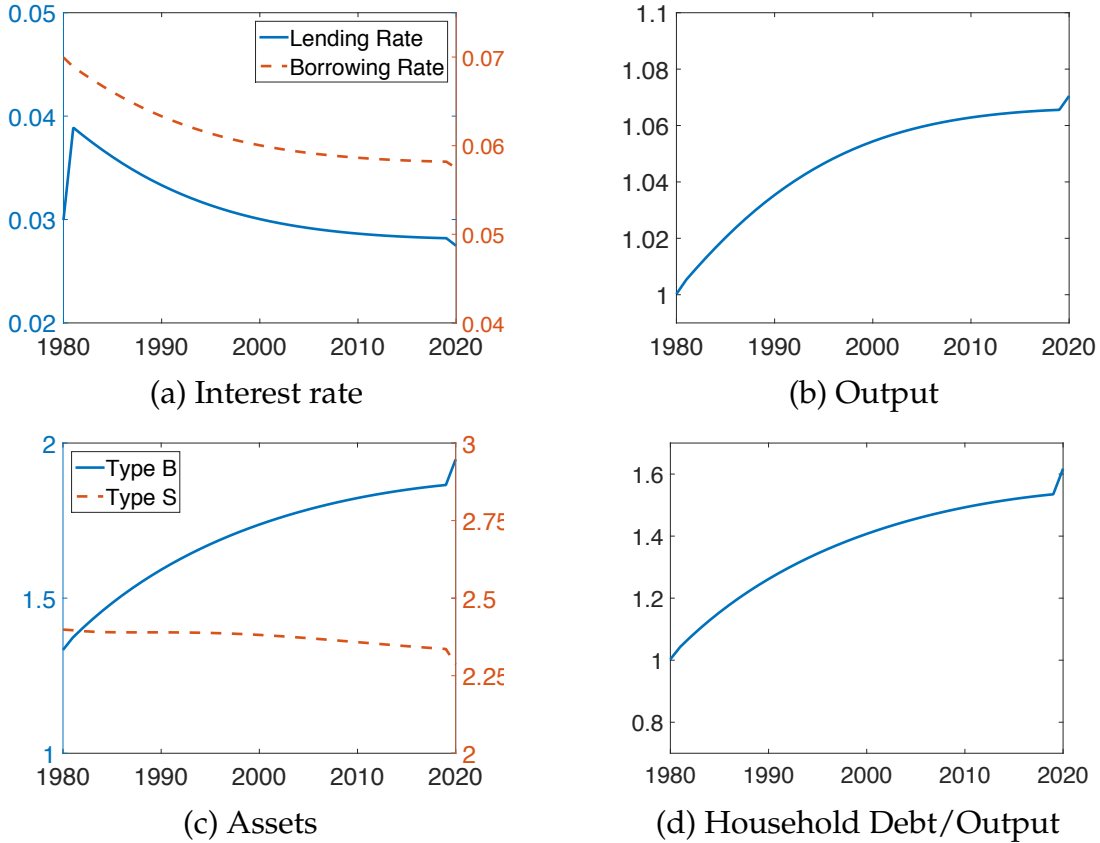
4.3   Transitions

Since it can take many years for an economy to converge to a new steady state, comparing two steady states may not be the best way to assess the impact of an increase in life expectancy in a four decade span. We show here that convergence indeed happen quite fast: by 2010 most of the increase in debt (around 90%) had already taken place. We also show that cyclical movements of productivity have played an important role in accommodating the slow growth in private debt observed during the early 80s and the subsequent speeding up during the 2000s.

---

[23]In our setting, the spread ϕ remains constant as we do not model the convenience yield of government bonds.

Figure 8: Transition Dynamics: Constant TFP

(a) Interest rate

(b) Output

(c) Assets

(d) Household Debt/Output

The computation of the transition presents several challenges. First, as there is a distribution of agents indexed by age and assets at the time the life expectancy increases, who is affected by the shock? We assume that all working-age agents experience in 1980 a decline in the survival probability to $\delta = 0.052$, while there is no change for retired agents. Second, as some agents were already involved in a banking contract, what happens with those contracts? We assume that after the shock all existing contracts are renegotiated to take into account the new survival probability. Third, what happens with the government budget? We assume that lump-sum transfers remain at the same absolute value as before the shock and that the government still follows a policy of maintaining the debt to output ratio constant and equal to 0.3, adjusting labor taxes correspondingly during the transition to maintain government's budget balanced. Finally, what happens with retirement payments? We assume that they do not change for those already retired in 1980.

In Figure 8 we show the transition dynamics when both life expectancy increases

and intermediation costs decline. In panel (a), we see that the spread (the differ-ence between lending and borrowing rates) falls drastically on impact and then in-creases slowly until reaching the new steady state. The lending rate converges non-monotonically because the capital stock is low with respect to its desired value when agents expect to live longer. Thus, the return on savings suddenly increases and then slowly converges to the new lower level as capital increases. The increase in capital induces a continuous increase in output (panel b) and an in the net worth of B-agents. Interestingly, the net worth of C-agents declines (panel c) in spite of the increase in capital because of the increase in leverage (panel d).

For convenience, in the panels we show the new steady state in the last period (in 2020) to get a sense of how complete the convergence is 40 years after the changes. In panel (d), for example, household debt increases from 1GDP in 1980 to about 1.54 by 2007, almost 85% of the difference between steady states.

Figure 9: Transition Dynamics: Observed TFP

(a) Interest rate

(b) Output

(c) Assets

(d) Household Debt/Output
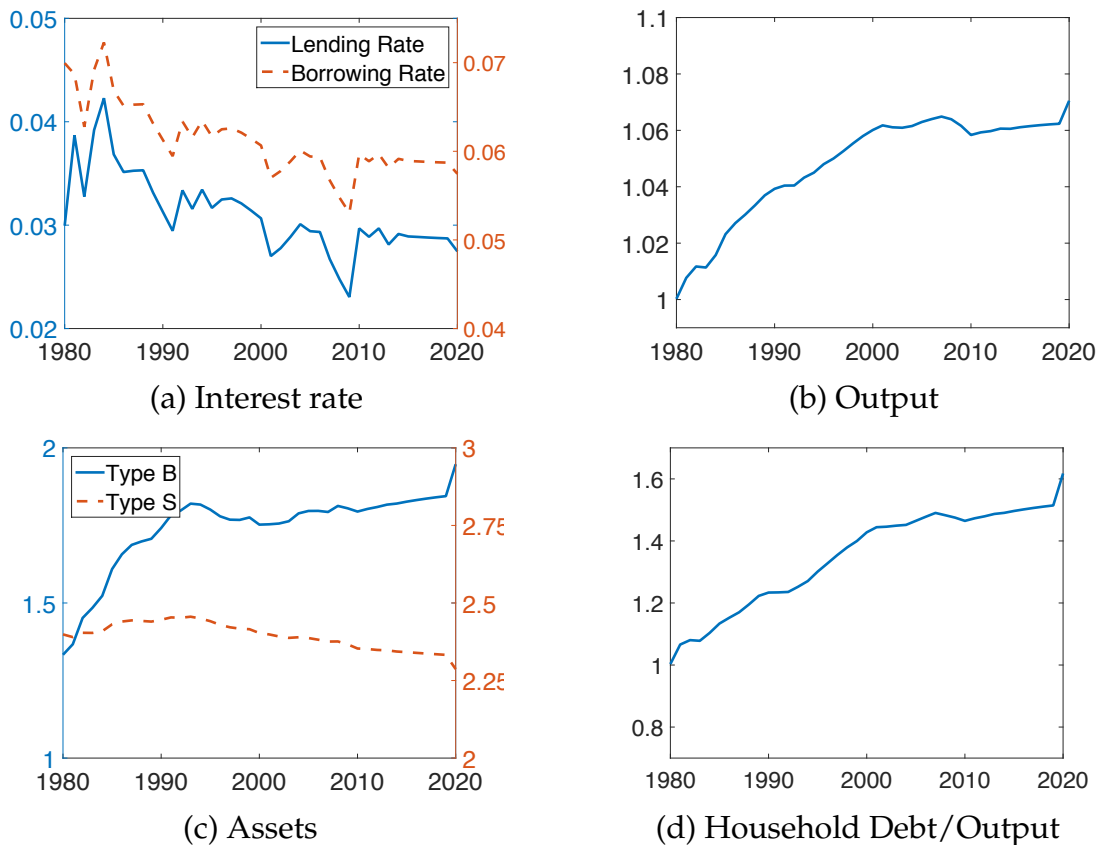
In Figure 9 we compute transitions using the actual path for TFP (measured by the Solow residual) instead of fixing $A = 1$, as in the previous simulation. This exercise is informative because it shows how the recessions in the early 80s and early 90's slow down the convergence during the 80s, and how debt speeds up in the second half of the 90s. Although the evolution of interest rates is mostly unaffected by TFP changes, output and private debt are slightly above the figures computed in Figure 8.[24]

Remark on the growth of shadow banking: Since our counterfactuals just compare scenarios with and without shadow banking, we are not required to discipline κ. It is outside the scope of this paper to introduce a distribution of κ across agents (capturing, for instance, heterogeneity on the ability of different agents to face a collapse in securitization or heterogeneity of search and informational costs required to operate using shadow banking) such that agents gradually move from traditional to shadow banking along the transition, as observed in the data. The speed at which shadow banking is adapted would discipline the distribution of κ.

# 5   On the Costs and Benefits of Shadow Banking

Our goal is to understand the rise and the benefits of shadow banking, thus abstracting from its potential cost in inducing a crisis. However, the great recession, characterized by a collapse in securitization and other instruments used to improve asset liquidation, is a reminder of how large these costs can be. To address this issue we appeal to the methodology proposed by Luttrell, Atkinson, and Rosenblum (2013), and later expanded by Ball (2014) and Fernald (2014). This approach compares the realized output with the potential output computed by the Congressional Budget Office (CBO), which implies that the great recession generated a loss, in present value, of 23% of 2007 GDP. In Figure 10, we depict this number as the difference between the dotted back line (potential GDP computed by the CBO) and the dashed red line (realized output) after 2007.[25]

---

[24]See Chen, Imrohoroglu, and Imrohoroglu (2006) and Fernandez, Imrohoroglu, and Tamayo (2018) for the effect of TFP in saving rates.

[25]For a detailed explanation of these calculations, see the staff report by Atkinson, Luttrell, and Rosenblum (2013). Our number is lower than the estimation from the previous literature, which ranges from 40% to 90% of 2007 GDP, mainly because the CBO has recently revised down potential output.

Figure 10: The Costs and Benefits of Shadow Banking
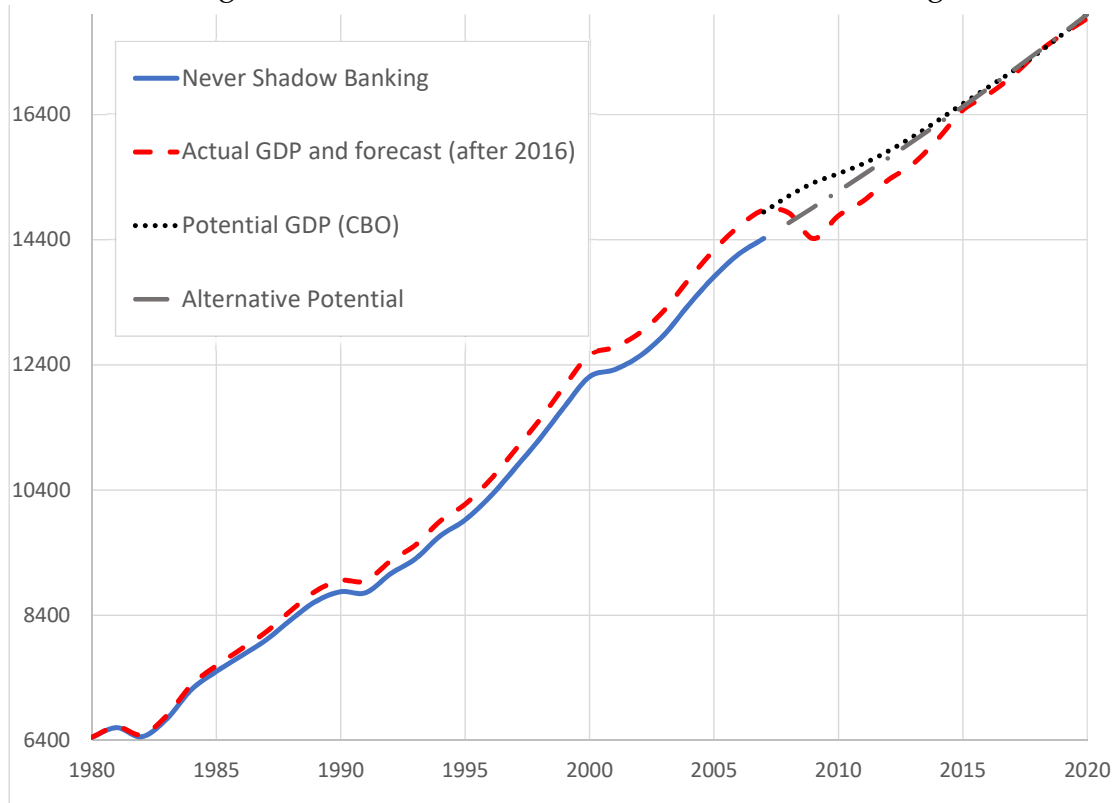
If securitization was the single responsible of the crisis, was it worth it? Was the contribution of shadow banking in the economy large enough to compensate this cost? A quick answer is provided by Table 1. Shadow banking generates a permanent increase in output level equivalent to 2.8% per year in the stationary equilibrium ($1.062 - 1.034$, according to the second and fourth columns of Table 1), which represents a present value of around 3.3GDP of 2007. This is, however, misleading. First, it assumes shadow banking is permanent and does not generate crises. Second, it overestimates the gains during the transition.

To make a more meaningful comparison of the benefits and cost of shadow banking surrounding the recent crisis, we compute a benchmark economy without shadow banks. This benchmark does not have the gains from lower intermediation costs, but it does not have the cost of a crisis either. For this counterfactual we assume $\kappa > \bar{\kappa}$, so that after the increase in life expectancy individuals keep choosing traditional banks and spreads remain at the 1980 level of 4%. This counterfactual is the blue solid line of Figure 10 until 2007 and the dashed-dotted grey line after 2007.

With our counterfactual we can now compute the gains from shadow banking before 2007 and its losses after 2007. The present value of the gap between the benchmark and the realized output from 1980 to 2007 represents 59% of 2007 GDP. The gap between the benchmark and the realized output from 2007 to 2020 represents 13.5% of 2007 GDP. Assuming that shadow banking was solely responsible of the crisis, the comparison between these two numbers still delivers a net gain from shadow banking of 45% of 2007 GDP.

Note that our estimated cost of shadow banking based on the counterfactual is lower than the estimated cost of 23% of 2007 GDP that, following the literature, is based on the CBO estimated potential output. The reason is that the initial level of the CBO potential output is the realized output before the crisis, which according to our model was initially high exactly because shadow banking was instrumental in increasing output. In other words, measuring the cost of a financial crisis using potential output is highly misleading as it ignores the forces that generated the output level before the crisis happens, confusing the ex-post cost of shadow banking with its ex-ante value.

6 Conclusions

The recent discussion, both in academic and policy circles, about the demand for safe assets and its macroeconomic effects has focused on the "savings glut" from foreign countries. At the same time, the recent discussion about shadow banking has focused on its pervasive role on triggering painful crises. In this paper we argue that these two discussions are intimately related. While the higher *foreign demand* for safe assets seems to have been accommodated by an increase in government debt, the higher *domestic demand* for safe assets triggered by an increase in life expectancy has pushed an endogenous increase in the supply of safe assets by the private sector, more specifically using securitization and shadow banking. We have explored quantitatively the individual roles of higher life expectancy and the rise of shadow banking in the accumulation of financial assets, private debt, output and welfare.

By calibrating our model to the eighties, we show that the domestic savings glut and the corresponding expansion of shadow banking accommodates well the large increase in asset accumulation and private credit experienced by the United States in the last four decades. We find that in the absence of shadow banking capital-output

ratio and output would have increased only half of what they did, while assets and credit would have not increased at all. Our approach allows us to compute a counterfactual without shadow banks. We have shown that the gains from operating with shadow banking from 1980 to 2007 were in the order of 60% of 2007 GDP. Further, even if we blame the great recession exclusively to the operation of shadow banks, its cost was in the order of 14% of 2007 GDP. In short, our model suggests that there were net gains to having shadow banks, even if it were true that they single handedly generated the recent crisis.

These results are relevant to the recent policy discussion about how to regulate the banking system. Although restricting the use of shadow banks or certain financial innovations, such as securitization, may have benefits in terms of reducing the likelihood and magnitude of financial crises, we show that it is also costly in terms of choking-off output. Even though our quantitative estimations are based on a streamlined model, and as such should be only taken as a "proof of concept" that the involved magnitudes are likely to be sizable, they highlight the relevance of measuring benefits and costs of shadow banking before implementing regulatory changes.

References

Alvarez, Fernando, Luigi Guiso, and Francesco Lippi. 2012. "Durable Consumption and Asset Management with Transaction and Observation Costs." *American Economic Review, Papers and Proceedings* 102 (5): 2272–2300.

Ameriks, John, Joseph Briggs, Andrew Caplin, Minjoon Lee, Matthew Shapiro, and Christopher Tonetti. 2019. "Older Americans Would Work Longer If Jobs Were Flexible." American Economic Journal: Macroeconomics. Forthcoming.

Ameriks, John, Andrew Caplin, Minjoon Lee, Matthew Shapiro, and Christopher Tonetti. 2014. "The Wealth of Wealthholders." Working Paper, Stanford University.

Atkinson, Tyler, David Luttrell, and Harvey Rosenblum. 2013, July. "How bad was it? The costs and consequences of the 2007–09 financial crisis." Technical Report 20, Federal Reserve Bank of Dallas.

Auclert, Adrien, Frederic Martenet, and Hannes Malmberg. 2019. "Wealth and Demographics in the 21st Century." 2019 meeting papers 952, Society for Economic Dynamics.

Ball, Laurence. 2014. "Long-Term Damage from the Great Recession in OECD Countries." NBER Working Paper 20185.

Bates, Thomas, Kathleen Kahle, and Rene Stulz. 2009. "Why Do U.S. Firms Hold So Much More Cash Than They Used To?" *Journal of Finance* 64:1985–2021.

Begenau, Juliane, and Tim Landvoigt. 2017. "Financial Regulation in a Quantitative Model of the Modern Banking System." Working Paper, Stanford University.

Bernheim, Douglas. 1991. "How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities." *Journal of Political Economy* 99:899–927.

Bertaut, Carol, Laurie Pounder DeMarco, Steven Kamin, and Ralph Tryon. 2012. "ABS inflows to the United States and the global financial crisis." *Journal of International Economics* 88 (2): 219–234.

Bloom, David E., David Canning, Richard K. Mansfield, and Michael Moore. 2007. "Demographic change, social security systems, and savings." *Journal of Monetary Economics* 54 (1): 92 – 114. Carnegie-Rochester Conference Series on Public Policy.

Bloom, David E., David Canning, and Michael Moore. 2014. "Optimal Retirement with Increasing Longevity." *Scandinavian Journal of Economics* 116 (3): 838–858.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang. 2017, October. "The Effects of Marriage-Related Taxes and Social Security Benefits." NBER Working Paper 23972.

Caballero, Ricardo. 2010. "A Caricature (Model) of the World Economy." Working Paper, MIT.

Caballero, Ricardo, Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2016. "Global Imbalances and Currency Wars at the ZLB." Working Paper, MIT.

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio. 2016. "Demographics and real interest rates: Inspecting the mechanism." *European Economic Review* 88 (C): 208–226.

Chen, Kaiji, Selahattin Imrohoroglu, and Ayse Imrohoroglu. 2006. "The Japanese Saving Rate." *American Economic Review* 96 (5): 1850–1858 (December).

Corbae, Dean, and Pablo D'Erasmo. 2018. "Capital Requirements in a Quantitative Model of Banking Industry Dynamics." Working Paper, Federal Reserve Bank of Philadelphia.

Costa, Dora L. 1998. "The Evolution of Retirement: Summary of a Research Project." *The American Economic Review* 88 (2): 232–236.

Dang, Tri Vi, Gary Gorton, Bengt Holmstrom, and Guillermo Ordonez. 2017. "Banks as Secret Keepers." *American Economic Review* 107:1005–1029.

Del Negro, Marco, Gauti Eggertsson, Andrea Ferrero, and Nobuhiro Kiyotaki. 2017a. "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities." *American Economic Review* 107 (3): 824–857 (March).

Del Negro, Marco, Domenico Giannone, Marc Giannoni, and Andrea Tambalotti. 2017b. "Safety, Liquidity, and the Natural Rate of Interest." *Brookings Papers on Economic Activity* 48 (1 (Spring): 235–316.

De Nardi, Mariacristina, Eric French, and John Jones. 2009. "Life Expectancy and Old Age Savings." *American Economic Review, Papers and Proceedings* 99 (2): 110–115.

———. 2010. "Why Do the Elderly Save? The Role of Medical Expenses." *Journal of Political Economy* 118:39–75.

———. 2015. "Savings After Retirement: A Survey." NBER Working Paper no. 21268.

Diamond, Douglas, and Philip Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91:401–419.

Diamond, Douglas, and Raghuram Rajan. 2001. "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109:287–327.

Diaz-Gimenez, Javier, Edward Prescott, Terry Fitzgerald, and Fernando Alvarez. 1992. "Banking in computable general equilibrium economies." *Journal of Economic Dynamics and Control* 16:533–559.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso. 2017. "Asset Market Participation and Portfolio Choice over the Life-Cycle." *The Journal of Finance* 72 (2): 705–750 (March).

Fagereng, Andreas, Martin Blomhoff Holm, Benjamin Moll, and Gisle Natvik. 2019. "Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains." Technical Report.

Farhi, Emmanuel, and Jean Tirole. 2017. "Shadow Banking and the Four Pillars of Traditional Financial Intermediation." Working Paper, Havard University.

Fernald, John. 2014. "Productivity and Potential Output Before, During and After the Great Recession." Federal Reserve Bank of San Francisco Working Paper 2014-15.

Fernandez, Andres, Ayse Imrohoroglu, and Cesar Tamayo. 2018. "Saving Rates in Latin America: A Neoclassical Perspective." 2018 meeting papers 1229, Society for Economic Dynamics.

Gale, William, and John Karl Scholz. 1994. "Intergenerational transfers and the accumulation of wealth." *Journal of Economic Perspectives* 8 (4): 145–160.

Gao, Xiaodan, Toni Whited, and Na Zhang. 2018. "Corporate Money Demand." Working paper, University of Michigan.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel. 2015. "Very Long-Run Discount Rates." *The Quarterly Journal of Economics* 130 (1): 1–53.

Gorton, Gary, and Guillermo Ordonez. 2014. "Collateral Crises." *American Economic Review* 104 (2): 343–378.

Gustman, Alan, and Thomas Steinmeier. 1999. "Effects of pensions on savings: Analysis with data from the Health and Retirement Study." *Carnegie-Rochester Conference Series on Public Policy*, Volume 50. Elsevier, 271–324.

Harris, Milton, Christian Opp, and Marcus Opp. 2014. "Higher Capital Requirements, Safer Banks? Macroprudential Regulation in a Competitive Financial System." Working Paper, Wharton School.

Imrohoroglu, Ayse, Selahattin Imrohoroglu, and Douglas H Joines. 1998. "The Effect of Tax-Favored Retirement Accounts on Capital Accumulation." *American Economic Review* 88 (4): 749–768 (September).

Imrohoroglu, Ayşe, and Kai Zhao. 2018. "The chinese saving rate: Long-term care risks, family insurance, and demographics." *Journal of Monetary Economics* 96 (C): 33–52.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti. 2013. "The Effects of the Saving and Banking Glut on the U.S. Economy." In *NBER International Seminar on Macroeconomics 2013*, NBER Chapters, 52–67. National Bureau of Economic Research, Inc.

———. 2015. "Credit Supply and the Housing Boom." Manuscript.

Kotlikoff, Laurence J., and Lawrence H. Summers. 1981. "The Role of Intergenerational Transfers in Aggregate Capital Accumulation." *Journal of Political Economy*, pp. 706–732.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120:233–267.

Krueger, Dirk, and Felix Kubler. 2005. "Pareto Improving Social Security when Financial Markets are Incomplete." *American Economic Review* 96:737–755.

Lockwood, Lee. 2012. "Bequest Motives and the Annuity Puzzle." *Review of Economic Dynamics* 15 (2): 706–732.

———. 2015. "Incidental Bequests: Bequest Motives and the Choice to Self-Insure Late-Life Risks." Manuscript.

Luttrell, David, Tyler Atkinson, and Harvey Rosenblum. 2013. "Assessing the costs and consequences of the 2007 - 09 financial crisis and its aftermath." *Economic Letter*, vol. 8 (Septiembre).

Mankiw, N. Gregory, and Stephen Zeldes. 1991. "The consumption of stockholders and nonstockholders." *Journal of Financial Economics* 29 (1): 97–112.

Mehra, Rajnish, Facundo Piguillem, and Edward Prescott. 2011. "Costly Financial Intermediation in Neoclassical Growth Theory." *Quantitative Economics* 2 (1): 1–36.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The equity premium: A puzzle." *Journal of Monetary Economics* 15 (2): 145–161 (March).

Moreira, Alan, and Alexi Savov. 2015. "The Macroeconomics of Shadow Banking." Working Paper, Yale University.

Ordonez, Guillermo. 2018a. "Confidence Banking and Strategic Default." *Journal of Monetary Economics*, vol. Forthcoming.

———. 2018b. "Sustainable Shadow Banking." *American Economic Journal: Macroeconomics* 10:33–56.

Philippon, Thomas. 2015. "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation." *American Economic Review* 105 (4): 1408–38.

Plantin, Guillaume. 2015. "Shadow Banking and Bank Capital Regulation." *Review of Financial Studies* 28 (1): 146–175.

Samuelson, Paul. 1958. "An exact consumption-loan model of interest with or without the social contrivance of money." *Journal of Political Economy* 66:467–482.

Scharfstein, David. 2018. "Presidential Address: Pension Policy and the Financial System." *Journal of Finance* 73:1463 – 1512.

Shourideh, Ali, and Maxim Troshkin. 2017. "Incentives and Efficiency of Pension Systems." Working Paper, Carnegie Mellon University.

Summers, Lawrence, and Christopher Carroll. 1987. "Why Is U.S. National Saving So Low?" *Brookings Papers on Economic Activity* 18 (2): 607–642.

Wolff, Edward. 2004. "Changes in household wealth in the 1980s and 1990s in the U.S." Levy Economics Institute Working Paper 407.

Appendix

A   Proof of Proposition 2:

We first characterize the best banking strategy in a steady state for a general interest rate r. This problem solves:

$$\max_{\{c_j, b_j\}} \left\{ \sum_{j=0}^{T} \beta^j \log c_j + \sum_{j=T+1}^{\infty} \beta^j (1-\delta)^{j-T-1}[(1-\delta)\log c_j + \delta\alpha \log b_j \right\}$$

$$s.t. \sum_{j=0}^{T} \frac{c_j}{(1+r)^j} + \sum_{j=T+1}^{\infty} \frac{c_j(1-\delta)^{j-T}}{(1+r)^j} + \sum_{j=T+1}^{\infty} \frac{b_j(1-\delta)^{j-T-1}\delta}{(1+r)^j} \leq v_0^B$$

Notice that the price of an annuity payment at age j is $P_j = \frac{1}{(1+r)^j}$ if the agent is alive at age $j \leq T$, $P_j = \frac{(1-\delta)^{j-T}}{(1+r)^j}$ if the agent is alive at age $j > T$ and $P_j = \frac{(1-\delta)^{j-T-1}\delta}{(1+r)^j}$ if the agent dies at age $j > T$. Thus, prices are present discounted values of the probabilities of each potential event contingent on age (and only on age).

The first order conditions for this problem generate:

$$c_{j+1} = \beta(1+r)c_j; \quad \forall j$$

$$b_j = \alpha c_j; \quad \forall j > T$$

These two equations imply:

$$c_j = \beta^j (1+r)^j c_0; \quad \forall j$$

$$b_j = \alpha\beta^j (1+r)^j c_0; \quad \forall j > T$$

Replacing the last two in the budget constraint we can find c_0, by solving:

$$\sum_{j=0}^{T} \frac{\beta^j (1+r)^j c_0}{(1+r)^j} + \sum_{j=T+1}^{\infty} \frac{\beta^j (1+r)^j c_0 (1-\delta)^{j-T}}{(1+r)^j} + \sum_{j=T+1}^{\infty} \frac{\alpha\beta^j (1+r)^j c_0 (1-\delta)^{j-T-1}\delta}{(1+r)^j} = v_0^B$$

Which gives as $c_0 = \bar{c}(\delta)v_0^B$, and then all consumptions are proportional to initial wealth, where

$$\bar{c}(\delta) = \frac{1-\beta}{1 - \beta^T + (1-\beta)\beta^T \theta_B(\delta)}. \tag{13}$$

49

We can simplify the characterization by splitting the problem in two parts: before and after retirement, which is useful in Lemma 2 where agents following strategy S change their pattern of consumption after retirement.

We guess and verify that the maximum utility after T can be written in recursive way as $\phi_B + \theta_B \log(w_T^B)$. For this to be true, the coefficients ϕ_B and θ_B must satisfy:

$$\phi_B + \theta_B \log(w) = \log(c) + \beta(1-\delta)[\phi_B + \theta_B \log(w')] + \beta\delta\alpha \log(b')$$

The problem after retirement solves:

$$\max_{c,w',b'} \{\log(c) + \beta(1-\delta)[\phi_B + \theta_B \log(w')] + \beta\delta\alpha \log(b')\}$$

$$s.t. \quad c + \frac{1-\delta}{1+r}w' + \frac{\delta}{1+r}b' \le w$$

which generates the first order conditions

$$w' = \beta(1+r)\theta_B c$$

$$b' = \beta(1+r)\alpha c$$

Substituting these in in the budget constraint we get that ,

$$c\left[1 + (1-\delta)\beta\theta_B + \delta\beta\alpha\right] = w \tag{14}$$

Now we guess that $c = \frac{w}{\theta_B}$ and verify it for

$$\theta_B(\delta) = \frac{1 + \beta\alpha\delta}{1 - \beta(1-\delta)},$$

confirming that the solutions are proportional to wealth.

Based on this solution, the maximum utility in steady state as a function of δ attainable by an agent who follows a banking strategy (B) that pays an interest r_s, where $s \in \{SB, TB\}$ is the indicator for whether the interest rate corresponds to shadow or traditional banking respectively, such that $r_{SB} > r_{TB}$, as shown in Proposition 1, can be expressed as:

$$U_B(\delta, r_s) = \sum_{j=0}^{T-1} \beta^j \log(c_j^B) + \beta^T[\phi_B + \theta_B \log(w_T^B)]$$

50

where

$$\theta_B(\delta) = \frac{1 + \beta\alpha\delta}{1 - \beta(1 - \delta)}$$

$$\phi_B(\delta, r_s) = \frac{(\theta_B(\delta) - 1)\log(\beta(1 + r_s)) - \log(\theta_B(\delta)) + \beta\alpha\delta[\log(\alpha) - \log(\theta_B(\delta))]}{1 - \beta(1 - \delta)}$$

$$c_j^B(\delta, r_s) = \bar{c}(\delta)\beta^j(1 + r_s)^j v_0^B$$

$$w_T^B(\delta, r_s) = \theta_B(\delta)\bar{c}(\delta)\beta^T(1 + r_s)^T v_0^B$$

where $\bar{c}$ is defined in equation (13) and v_0^B in equation (2).

Define $\Delta_B(\delta) = [U_B(\delta, r_{SB}) - \kappa] - U_B(\delta, r_{TB})$. Lemma 1 below shows that, as long as α is not too high, $\frac{\partial \Delta_B(\delta)}{\partial \delta} < 0$, i.e., the utility difference of participating in shadow banking is increasing in life expectancy (decreasing in δ) for all $\delta > 0$.

Lemma 1 If $\alpha < \frac{1}{1-\beta}$ then $\frac{\partial \Delta_B(\delta)}{\partial \delta} < 0, \quad \forall \delta > 0$.

Proof Using the property of logarithmic functions,

$$\Delta_B(\delta) = \sum_{j=0}^{T-1} \beta^j \log\left[\frac{(1 + r_{SB})^j}{(1 + r_{TB})^j}\right] + \beta^T \left[\phi_B(\delta, r_{SB}) - \phi_B(\delta, r_{TB}) + \theta_B(\delta)\log\left(\frac{w_T^B(r_{SB})}{w_T^B(r_{TB})}\right)\right] - \kappa$$

where $\phi_B(\delta, r_{SB}) - \phi_B(\delta, r_{TB}) = \widehat{\theta}_B(\delta)\log\left(\frac{1 + r_{SB}}{1 + r_{TB}}\right)$, defining

$$\widehat{\theta}_B(\delta) = \frac{\beta(1 + \delta(\alpha - 1))}{[1 - \beta(1 - \delta)]^2}.$$

Then, we can rewrite the new benefit of shadow banking as

$$\Delta_B(\delta) = \sum_{j=0}^{T-1} \beta^j \log\left[\frac{1 + r_{SB}}{1 + r_{TB}}\right]^j + \beta^T \log\left(\frac{1 + r_{SB}}{1 + r_{TB}}\right)\left[\widehat{\theta}_B(\delta) + \theta_B(\delta)\log\left(\frac{1 + r_{SB}}{1 + r_{TB}}\right)^{T-1}\right] - \kappa$$

Taking derivatives with respect to δ,

$$\frac{\partial \Delta_B(\delta)}{\partial \delta} = \beta^T \log\left(\frac{1 + r_{SB}}{1 + r_{TB}}\right)\left[\frac{\partial \widehat{\theta}_B(\delta)}{\partial \delta} + \frac{\partial \theta_B(\delta)}{\partial \delta}\log\left(\frac{1 + r_{SB}}{1 + r_{TB}}\right)^{T-1}\right]$$

where

$$\frac{\partial \widehat{\theta}_B(\delta)}{\partial \delta} = \frac{\beta[(\alpha - 1)(1 - \beta(1 + \delta)) - 2\beta]}{[1 - \beta(1 - \delta)]^3}$$

51

and

$$\frac{\partial \theta_B(\delta)}{\partial \delta} = \frac{\beta[\alpha(1-\beta)-1]}{[1-\beta(1-\delta)]^2}.$$

Notice that $\frac{\partial \widehat{\theta}_B(\delta)}{\partial \delta} < 0$ if and only if $\alpha < \frac{(1-\beta\delta)+\beta}{(1-\beta\delta)-\beta}$ and $\frac{\partial \theta_B(\delta)}{\partial \delta} < 0$ if and only if $\alpha < \frac{1}{1-\beta}$. Since the first condition is always satisfied when the second condition is satisfied, then the sufficient condition for $\frac{\partial \Delta_B(\delta)}{\partial \delta}$ is that $\alpha < \frac{1}{1-\beta}$.

Notice that the conditions for an interior δ^* are $\Delta_B(0) > 0$, this is

$$\kappa < \sum_{j=0}^{T-1} \beta^j \log \left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j + \beta^T \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right) \left[\frac{\beta}{(1-\beta)^2} + \frac{1}{1-\beta} \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right]$$

and $\Delta_B(1) < 0$, this is

$$\kappa > \sum_{j=0}^{T-1} \beta^j \log \left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j + \beta^T \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right) \left[\beta\alpha + (1+\beta\alpha) \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right]$$

which is feasible when $1 - \alpha(1-\beta) > 0$, or $\alpha < \frac{1}{1-\beta}$. Q.E.D.

Since $U_B(\delta, r_{SB}) - U_B(\delta, r_{TB})$ is independent of κ and $\Delta_B(\delta)$ is just linear in κ it is straightforward from Lemma 1 that, given κ there is a single $\delta^* \in (0,1)$ such that $\Delta_B(\delta^*) = 0$, where $\delta^* = 0$ if $\Delta_B(0) < 0$ and $\delta^* = 1$ if $\Delta_B(1) > 0$. Furthermore, δ^* weakly decreases in κ (strictly except at the corners, where κ is so high that $\delta^* = 0$ or so low that $\delta^* = 1$).

Finally, computing $\frac{\partial \Delta(\delta)}{\partial \alpha}$ it is easy to see that

$$\frac{\partial \Delta_B(\delta)}{\partial \alpha} = \beta^T \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right) \left[\frac{\partial \widehat{\theta}_B(\delta)}{\partial \alpha} + \frac{\partial \theta_B(\delta)}{\partial \alpha} \log \left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right] > 0.$$

This derivative is positive because $\frac{\partial \widehat{\theta}_B(\delta)}{\partial \alpha} = \frac{\beta\delta}{[1-\beta(1-\delta)]^2} > 0$ and $\frac{\partial \theta_B(\delta)}{\partial \alpha} = \frac{\beta\delta}{1-\beta(1-\delta)} > 0$. This implies that, fixing δ and κ, δ^* is weakly increasing in α (strictly increasing except at the corners). QED.

B An interpretation of κ:

Even though we have introduced the cost κ on shadow banking in a reduced form way, here we show that we can write

$$\kappa = -\beta^T \frac{p \log(\zeta)}{1 - \beta(1-\delta)} > 0$$

where $p < 1$ is the defined as the yearly probability that the bank cannot pay as promised, in which case the agent just consumes a fraction $\zeta < 1$ of the promised consumption. Notice that we assume that securitization can enter into a crisis ($p > 0$), while traditional banking cannot (see Gorton and Ordonez (2014), for microfoundations of such a crisis due to information opacity and lack of government explicit support). Thus κ can be interpreted as the net cost of shadow banking. Furthermore, the lower the recovery in case of a crisis (lower ζ), the higher the cost.

This expression comes from extending the recursive formulation of the utility conditional on retirement as

$$
\begin{aligned}
\phi_B + \theta_B \log(w) &= [(1-p)\log(c) + p\log(\zeta c)] + \beta(1-\delta)[\phi_B + \theta_B \log(w')] + \beta\delta\alpha\log(b') \\
&= [\log(c) + \beta(1-\delta)[\phi_B + \theta_B\log(w')] + \beta\delta\alpha\log(b')] + p\log(\zeta)
\end{aligned}
$$

which adds a constant compared to the previous specification without crises (this is, with $p = 0$ or $\zeta = 1$). As this does not affect the first order conditions, $\theta_B(\delta)$ remains unchanged, but the constant term is affected as

$$
\phi_B(\delta, r_s, p, \zeta) = \phi_B(\delta, r_s) + \frac{p\log(\zeta)}{1 - \beta(1-\delta)}
$$

and then $\phi_B(\delta, r_{SB}, p, \zeta) - \phi_B(\delta, r_{TB}) = \widehat{\theta}_B(\delta) \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right) + \frac{p\log(\zeta)}{1-\beta(1-\delta)}$.

Therefore, we can rewrite the benefit of shadow banking as

$$
\Delta_B(\delta) = \sum_{j=0}^{T-1} \beta^j \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j + \beta^T \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)\left[\widehat{\theta}_B(\delta) + \theta_B(\delta)\log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right] - \underbrace{\left[-\beta^T \frac{p\log(\zeta)}{1-\beta(1-\delta)}\right]}_{\kappa}
$$

The rest of the analysis follows, with the only exception that κ in this case also depend on δ (as δ declines the cost of shadow banking also increases). However, the adjusted condition for an interior δ^* are $\Delta_B(0) > 0$, this is

$$
-\beta^T p\log(\zeta) < (1-\beta)\left[\sum_{j=0}^{T-1} \beta^j \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j + \beta^T \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)\left[\frac{\beta}{(1-\beta)^2} + \frac{1}{1-\beta}\log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right]\right]
$$

and $\Delta_B(1) < 0$, this is

$$
-\beta^T p\log(\zeta) > \sum_{j=0}^{T-1} \beta^j \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j + \beta^T \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)\left[\beta\alpha + (1+\beta\alpha)\log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right]
$$

which is feasible when $\alpha \underbrace{\left(1 + \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right)}_{>1} < \frac{1}{1-\beta} - \underbrace{\frac{\sum_{j=0}^{T-1} \beta^j \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j}{\beta^T \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]}}_{>0}$.

53

This condition is more stringent than with the reduced form κ, but the insight is the same, as δ^* is well defined when agents have relatively low bequest motives, who are the agents who self-select into banking contracts. Also, because in our calibration $\alpha = 0$ and β is close to 1, the condition is satisfied in the quantitative exercise.

C Proof of Proposition 3:

Take a steady state with prices (r, r_e, y_0), tax rate τ, and inheritance $\bar{b}$ to any measure zero individual. Let $U_B(\alpha)$ and $U_C(\alpha)$ represent the maximum attainable utility of an agent of measure zero in this economy who follows strategy B (banking) or C (capital markets) respectively as a function of α. Define $\Delta(\alpha) = U_C(\alpha) - U_B(\alpha)$. Lemma 2 below shows that, as long as δ is not too small, $\frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$, i.e., the utility difference is increasing in the bequest motive, for all $\alpha \geq 0$.

Lemma 2 *If* $\frac{1+r_e}{1+r} > \beta \left[\frac{1-\beta(1-\delta)}{\beta\delta} \right]^{1-\beta(1-\delta)}$ *then* $\frac{\partial \Delta(\alpha)}{\partial \alpha} \geq 0, \quad \forall \alpha > 0.$

Proof The maximum utility as a function of α attainable by an agent who follows a banking strategy (B), taking as given the parameters of the economy, can be expressed as:

$$U_B(\alpha) = \sum_{j=0}^{T-1} \beta^j \log(c_j^B) + \beta^T [\phi_B(\alpha) + \theta_B(\alpha) \log(w_T^B)]$$

where

$$\theta_B(\alpha) = \frac{1 + \beta\alpha\delta}{1 - \beta(1 - \delta)}$$

$$\phi_B(\alpha) = \frac{(\theta_B(\alpha) - 1) \log(\beta(1 + r)) - \log(\theta_B(\alpha)) + \beta\alpha\delta[\log(\alpha) - \log(\theta_B(\alpha))]}{1 - \beta(1 - \delta)}$$

$$c_j^B = \bar{c}(\alpha)\beta^j(1 + r)^j v_0^B$$

$$w_T^B = \theta_B(\alpha)\bar{c}(\alpha)\beta^T(1 + r)^T v_0^B$$

where $\bar{c}(\alpha) = \frac{1-\beta}{1-\beta^T+(1-\beta)\beta^T\theta_B(\alpha)}$ and v_0^B is defined in equation (2).

Similarly, the maximum utility as a type α who saves in capital markets (C) is

$$U_C(\alpha) = \sum_{j=0}^{T-1} \beta^j \log(c_j^C) + \beta^T [\phi_C(\alpha) + \theta_C(\alpha) \log(w_T^C)]$$

54

where

$$\theta_C(\alpha) = \frac{1 + \beta\alpha\delta}{1 - \beta(1-\delta)}$$

$$\phi_C(\alpha) = \frac{(\theta_C(\alpha) - 1)\log(1 + r_e) + (\theta_C(\alpha) - 1)\log(\theta_C(\alpha) - 1) - \theta_C(\alpha)\log(\theta_C(\alpha))}{1 - \beta(1-\delta)}$$

$$c_j^C = \bar{c}(\alpha)\beta^j(1 + r_e)^j v_0^C$$

$$w_T^C = \theta_C(\alpha)\bar{c}(\alpha)\beta^T(1 + r_e)^T v_0^C$$

Since $\theta_C(\alpha) = \theta_B(\alpha) = \theta(\alpha)$, using the properties of the logarithm function:

$$\Delta(\alpha) = \sum_{j=0}^{T-1} \beta^j \log\left[\frac{(1 + r_e)^j v_0^C}{(1 + r)^j v_0^B}\right] + \beta^T \left[\phi_C(\alpha) - \phi_B(\alpha) + \theta(\alpha)\log\left(\frac{w_T^C}{w_T^B}\right)\right] \quad (15)$$

Because the first term is independent of α it follows that

$$\frac{\partial \Delta(\alpha)}{\partial \alpha} = \beta^T \frac{\partial(\phi_C(\alpha) - \phi_B(\alpha))}{\partial \alpha} + \beta^T \theta'(\alpha)\log\left(\frac{w_T^C}{w_T^B}\right) \quad (16)$$

where $\theta'(\alpha) = \frac{\beta\delta}{1-\beta(1-\delta)}$ which does not depend on α.

$$\frac{w_T^C}{w_T^B} = \frac{(1 + r_e)^T v_0^C}{(1 + r)^T v_0^b} = \frac{\sum_{j=0}^{T-1} \frac{(1-\tau)y_0(1+\gamma)^j}{(1+r_e)^{j-T}} + \frac{\bar{b}}{(1+r_e)^{T_I-T}}}{\sum_{j=0}^{T-1} \frac{(1-\tau)y_0(1+\gamma)^j}{(1+r)^{j-T}} + \frac{\bar{b}}{(1+r)^{T_I-T}}} > 1$$

since $r_e > r$, $j < T$ and $T > T_I$. This implies the second term in (16) is positive, i.e., $\beta^T \theta'(\alpha)\log(\frac{w_T^C}{w_T^B}) > 0$

To prove $\frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$, we proceed in three steps showing that:

a) $\lim_{\alpha \to 0} \frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$;

b) $\frac{\partial^2 \Delta(\alpha)}{\partial \alpha^2} < 0$;

c) $\lim_{\alpha \to +\infty} \frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$

Simple algebra yields

$$\frac{\partial(\phi_C(\alpha) - \phi_B(\alpha))}{\partial \alpha} = \frac{\theta'(\alpha)}{1 - \beta(1-\delta)}\left[\log\left(\frac{(1 + r_e)}{(1 + r)\beta}\right) + \log\left(\frac{\theta(\alpha) - 1}{\alpha}\right) - \beta(1 - \delta)\log(\frac{\theta(\alpha)}{\alpha})\right] \quad (17)$$

55

From (17) it is readily seen that $\lim_{\alpha \to 0} \frac{\partial(\phi_C(\alpha) - \phi_B(\alpha))}{\partial \alpha} \to +\infty$. This follows since the last term tends to $+\infty$ and all the other terms are bounded. This coupled with the fact that $\beta^T \theta'(\alpha) \log(\frac{w_T^C}{w_T^B}) > 0$ proves that $\lim_{\alpha \to 0} \frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$.

The second derivative $\frac{\partial^2 \Delta(\alpha)}{\partial \alpha^2} < 0$ is negative by direct differentiation,

$$\frac{\partial^2 \Delta(\alpha)}{\partial \alpha^2} = \frac{-\beta^{T+1}\delta(1-\delta)}{\alpha(1 - \beta(1-\delta))(1 + (\alpha-1)\delta)(1 + \alpha\beta\delta)} < 0$$

since the denominator is always positive and the numerator is negative.

Finally it can be shown that $\lim_{\alpha \to \infty} \frac{\partial \Delta(\alpha)}{\partial \alpha} > 0$ under the condition stated in the theorem. Notice that (taking the limit of (17) when $\alpha \to \infty$) equation (16) is positive if and only if

$$\frac{1}{1 - \beta(1-\delta)} \log\left(\frac{(1+r_e)}{(1+r)\beta}\right) + \log\left(\theta'(\alpha)\right) + \log\left[\frac{(1+r_e)^T v_0^C}{(1+r)^T v_0^B}\right]$$

The last term in the above expression has already been shown to be positive. Thus a sufficient condition for this inequality is

$$\frac{1}{1 - \beta(1-\delta)} \log\left(\frac{(1+r_e)}{(1+r)\beta}\right) + \log\left(\frac{\beta\delta}{1 - \beta(1-\delta)}\right) > 0$$

This inequality can be written as

$$\frac{1+r_e}{1+r} > \beta \left[\frac{1 - \beta(1-\delta)}{\beta\delta}\right]^{1-\beta(1-\delta)}$$

Since a), b), and c) are satisfied, it follows that $\frac{\partial \Delta(\alpha)}{\partial \alpha} > 0, \forall \alpha \geq 0.$ Q.E.D.

On the one extreme, if $r_e = r$, insurance is free and all agents would prefer to follow strategy B. Thus, $\Delta(\alpha) < 0, \forall \alpha \geq 0$. On the other extreme, as $r_e - r \to +\infty$, the returns from self-insurance are so large that $\Delta(\alpha) > 0 \ \forall \alpha \geq 0$. Because $\Delta(\alpha, \phi)$ is continuous in ϕ it follows that there exist $\underline{\phi}$ and $\bar{\phi}$ with $\underline{\phi} < \bar{\phi}$ such that there is a unique $\alpha^*(\delta)$ for which $\Delta(\alpha^*) = 0$. Then the Lemma 2 that we prove below delivers the existence and uniqueness of the threshold $\alpha^*(\delta)$. QED