Retirement in the Shadow (Banking)[*]

Guillermo Ordoñez[†]   Facundo Piguillem[‡]

January 2019

Abstract

The U.S. economy has recently experienced a large increase in life expectancy and in securitization that fueled shadow banking activities. We argue that these two phenomena are intimately related. Agents rely on financial intermediaries to insure consumption during their uncertain life spans after retirement. When they expect to live longer, they rely more heavily on financial instruments and intermediaries that are riskier but offer better insurance terms, such as securitization and shadow banks. We calibrate the model to replicate the level of financial intermediation in 1980, introduce the observed change in life expectancy and show that the demographic transition is critical in accounting for the boom in both shadow banking and credit that preceded the recent U.S. financial crisis. We compare the U.S. experience with a counterfactual without shadow banks and show that they may have contributed to accumulate around 0.6GDP to output, four times larger than the estimated costs of the crisis.

JEL classification: E21. E44.

[†]University of Pennsylvania and NBER (e-mail: ordonez@econ.upenn.edu).
[‡]EIEF and CEPR (e-mail: facundo.piguillem@gmail.com).

1   Introduction

From the 1980s to the great recession, the U.S. economy experienced a steep increase in inter-mediated credit, with household debt growing from 1GDP to 1.7GDP. Due to the magnitude of the subsequent financial crisis, policymakers and scholars have 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 in triggering crises and by denying its potential benefits, these explanations tend to deem the boom to be 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: the de-mographic transition characterized by a longer life. In just three decades, there was a dramatic increase in life expectancy conditional on retirement, from 77 years to around 83 years. This is unique to this time frame as the increase in life expectancy in previous decades was mostly driven by a dramatic decline in child mortality.[1]

This demographic change induced an increase in the demand for precautionary savings, which we argue opened the doors for new and more efficient ways to supply insurance – se-curitization and the so-called shadow banking. Saving for retirement is indeed one of the most important drivers of financial intermediation, as retirees hold a large fraction of total wealth. 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 that for house-holds 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.

The first panel of Figure 1 confirms to the naked eye that the massive increase in credit over GDP mirrors in magnitude the increase of pensions over GDP, which have increased from 0.6GDP to 1.1GDP since the 1970s. The second panel shows that shadow banking was critical in accommodating the increase. The portfolio composition of pension funds has changed from 70% corporate equity and debt securities and no mutual fund shares (a prominent conduit of shadow banking activities) to 35% of each.

In this paper, we show that i) this *domestic savings glut* can account for most of the observed credit boom; ii) shadow banking was instrumental in accommodating the larger demand for

---

[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.

Figure 1: Evolution of Pensions and the Role of Shadow Banking
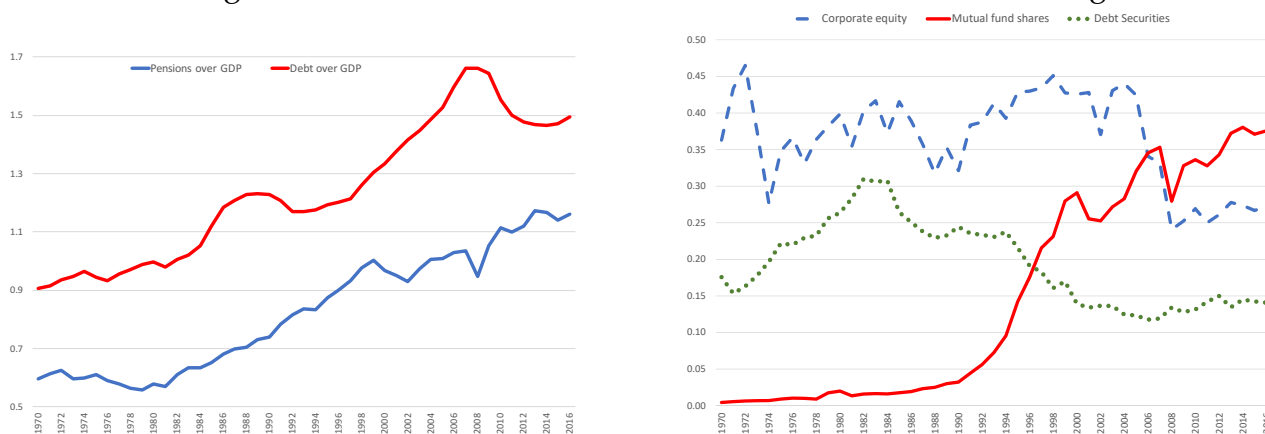
insurance, and it did so by substantially resorting to securitization and decreasing the financial sector's liquidity cost; and iii) even if we assume that the great recession was entirely caused by shadow banking operations, the benefits prior to the crisis were an order of magnitude larger than the 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*. We propose an overlapping generations model with heterogeneity in the bequest motives of individuals that allows for the coexistence of lenders and borrowers. Individuals with high-bequest motives save for retirement by buying capital, partly borrowing from individuals with low-bequest motives, who save for retirement by depositing their funds with financial intermediaries.[2]

All credit is managed through financial intermediaries, which channel funds from depositors to borrowers while guaranteeing depositors that their retirement savings are safe. The cost of the first activity, which we denote *operation cost*, is the cost of finding the best available investment opportunities to allocate the funds and includes the process of finding productive opportunities, monitoring the management of projects and administering payments. The cost of the second activity, 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 fraction of depositors larger than that expected to withdraw funds (those that are at their retirement age) choose to withdraw their funds in advance.

There are two types of financial intermediaries: traditional banks and shadow banks. The difference between them is that shadow banks face less constraints to use securitization, and

---

[2]The relevance of bequest motives to understanding annuities has been discussed by Bernheim (1991) and Lockwood (2012), among others.

therefore lower 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 for the assets being securitized. This technology exploits the diversification implicit in the pooling by targeting different investors, dividing the pool into tranches and in making transactions complex and opaque to discourage asymmetric information concerns among counterparties (see Ordonez (2018a) for a discussion about these elements). By operating at lower liquidity costs then, shadow banks can offer higher interest rates for deposits. However, the use of securitization comes at a cost in terms of fragility (sudden dry up of liquidity) inherent to the use of opaque operations. [3]

But how relevant were shadow banking operations in the United States to reduce liquidity costs? To address this question, *the second stage is empirical*. We show that the cost of intermediation, measured by the spread between lending and deposit rates, declined from a stable level of 4% in 1980 to around 3% before the recent financial crisis. We construct a measure of liquidity costs and show that the decline in intermediation costs can be explained almost completely by a decline in liquidity costs. This finding is consistent with Philippon (2015), who shows that operation costs have been constant in the financial sector for almost a century.

We then decompose the decline of liquidity costs and show that it can be mostly explained by the expansion of shadow banking. First, securitization improves the liquidity of productive risky loans, such that they can be traded with similar speed and discount as government bonds that are backed instead by unproductive safe taxation. 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 such as government bonds. These two forces allow for banks to more efficiently channel resources to the most productive opportunities.

But is this reduction of liquidity costs induced by shadow banking quantitatively consistent with the changes in volumes and prices of intermediation observed in the United States 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*.

---

[3]see Gorton and Ordonez (2014) for a microfoundation about sudden changes in information regimes related to securities.

Under certain conditions, for a given returns differential, as the life expectancy increases so does the present value of the gains to "depositing" in shadow banks. Intuitively, a higher life expectancy triggers an appetite for yields, and shadow banks can fulfill that appetite. In essence, an increase in life expectancy increases the demand for safe assets, forcing lower returns. This decrease in returns, together with the longer life expectancy, increases the benefits to searching for higher yields and the likelihood that shadow banking prospers, increasing the supply of safe assets and counteracting the decline in returns.

We calibrate the economy to 1980 and input the change in life expectancy and intermediation 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. 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, absent demographic changes, steady state output would have grown by only half the amount as with both forces combined, as the risk-free rate would have substantially increased. These results highlight the importance of first understanding the determinants of financial markets to then assessing their impact on aggregate dynamics.

To focus on our contribution, our model abstracts from the possibility that shadow banking collapses on path, but yet we discuss our initial question, one that has attracted fierce debate in policy and regulatory circles: did the United States win or lose from the operation of shadow banks? This justifies our *last, counterfactual, stage*. We construct a hypothetical economy without shadow banks and compare it with the realized economy in the U.S., with shadow banks and with a crisis that is completely attributed to the existence of shadow banks.

We find that, from 1980 to 2007, the existence 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.[4] Thus, even in the extreme case of blaming the crisis and its cost entirely on shadow banking activities, the economy gained almost half of 2007 GDP by the operation of shadow banking since the 1980s.

Related Literature: We contribute to the recent academic and policy discussion 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 crises and making financial systems fragile, much less is known about potential positive macroeconomic effects. As in our paper, Moreira and Savov (2015) highlight that shadow banking improves liquidity provision during booms and enhances growth, but at the cost of increasing fragility. In their case, shadow banking expands during periods of low uncertainty in the economy and collapses when uncertainty increases. In contrast, we take a longer perspective and study the role of higher retirement saving needs and a higher demand for safe assets in boosting the use of shadow banks. Even though we focus

---

[4]Following the literature and computing the costs by comparing realized output with potential GDP constructed by the U.S. Congressional Budget Office, instead of with our benchmark, the cost of the crisis is 23% of 2007 GDP.

on the positive macroeconomic effects of the run-up of shadow banking, not on its demise, we are able to provide an estimate of the net gains of shadow banking, even if we assign the whole blame for the crisis to its operation.

Similarly, we do not discuss the optimal regulation of shadow banking, but provide a quantitative assessment of its benefits and costs if the crisis were completely triggered by its existence. For a quantitative macroeconomic model of optimal regulation of shadow banking, see Begenau and Landvoigt (2017).[5]

In contrast to a rich literature (such as Caballero (2010), Caballero, Farhi, and Gourinchas (2016) and Carvalho, Ferrero, and Nechio (2016)) that argues that the increase in the demand for safe assets may have originated with foreign saving needs (the "global savings glut" hypothesis), in this paper we focus on the increase in the demand for safe assets coming from U.S. residents' higher needs for retirement savings (a "domestic savings glut"). Interestingly, a large part of the savings glut 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 not only that these forces are substantial quantitatively but also that a calibrated model can account for most of these changes.

The paper contributes more generally to the discussion on the relevance of savings for retirement to investment, output, and interest rates in macroeconomics when allowing explicitly for financial intermediation, as in Mehra, Piguillem, and Prescott (2011). We extend their environment by making the financial sector, in particular the roles of traditional and shadow banking, endogenous.

Our work captures in reduced form the same forces present in many microfoundations showing how shadow banks are able to reduce liquidity costs. This allows us to write a parsimonious model suitable to perform the macro quantitative analysis. Gorton and Ordonez (2014) show that securitization through pooling and tranching, the tool most used for shadow banking activities, reduced the incentives for information acquisition and allowed risky assets to be combined and traded as safe assets, providing "safety" at lower costs. This justifies our assumption of securitization reducing liquidity costs at a cost.

Similarly, we have refrain from the intricacies of adding regulatory consideration, but capturing them in reduced form that we can calibrate. Ordonez (2018b), for example, highlights that shadow banking is beneficial because it allows an escape from blunt regulations at the cost

---

[5]Other papers that focus on the interactions between regulation and shadow banking are Harris, Opp, and Opp (2014) and Plantin (2015).

of excessive risk-taking. Farhi and Tirole (2017) discuss how traditional banking is sustained on complementarities between costly public supervision and beneficial public liquidity guarantees, and how regulation (taxes and subsidies, ring fencing, etc.) can accommodate these forces to avoid a migration toward shadow banking.

Next we introduce a macroeconomic model with savings for retirement and financial intermediation, calibrate it and decompose the effects of retirement needs and shadow banking on welfare, output and the accumulation of assets.

2 Model

2.1 Environment

We study an overlapping generations economy populated by households that work in a competitive productive sector, save for retirement through financial intermediaries and are taxed by the government.

2.1.1 Households

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 can work an inelastic amount of hours without 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, which provides a utility $\alpha \geq 0$ (in units of consumption) per unit of bequest. Agents are heterogeneous in the intensity of their bequest motive, $\alpha \sim m(\alpha)$. Denote the consumption of an age-j agent at calendar time t by $c_{t,j}$. Assuming logarithmic preferences and a discount factor β, the utility present value of a agent that is born at a calendar period t is

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

Some remarks about this specification of preferences are in order. First, as is clear from equation (1), we assume the "joy-of-giving" type of bequest motive. This motive may capture, however, other forces. De Nardi, French, and Jones (2010) and De Nardi, French, and Jones (2015), for instance, show that agents save after retirement as a precaution against medical ex-

6

penses. As health is a normal good, the joy-of-giving specification also delivers this concern in a simple way. Thus, the reader must interpret the parameter α as capturing both precautionary savings against large potential health shocks in old age and pure bequest motives. As pointed out by De Nardi, French, and Jones (2015), it is extremely difficult, if not impossible, to properly disentangle the contribution of each effect.[6] Besides being instrumental in simplifying the solution of the model, this specification is also useful to capture non-trivial effects of changes in the age structure over savings.[7]

We have also assumed an exogenous retirement age. This is a simplifying assumption that still resembles the observed pattern of retirement in the U.S. 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 life after 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) that induces an earlier retirement. The final effect of a higher life expectancy in retirement age is then ambiguous. Costa (1998) indeed shows that the retirement age in the U.S., as in many other countries, has been continuously decreasing over the last 100 years, which points to the dominance of the income-wealth effect. As a result, with this assumption we would be underestimating the effect of aging on savings.[8]

Individuals have three sources of income. First, each agent born in period t receives labor income $y_{t,j}$ for the labor provided at age j during the first T years of her life (working age). Second, we assume that the bequest $b_{t,j}$ that agents leave upon death at age j is equally 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, individuals may receive pension transfers P^i_{t+j} from Social Security every period after retirement. Denoting agent i's saving returns by r^i_t and assuming a labor income tax τ, the agent i' that was born at t has a consolidated total wealth at birth of [9]

---

[6]See also Lockwood (2015) for an attempt to identify each component.

[7]For instance, if we had assumed 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. This would be at odds with the empirical evidence, as discussed by De Nardi, French, and Jones (2009).

[8]See also Bloom et al. (2007).

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

$$v_0^i = \sum_{j=0}^{T-1} \frac{(1-\tau)y_j}{(1+r^i)^j} + \frac{\bar{b}}{(1+r^i)^{T_I}} + \frac{(1+r^i)}{r^i+\delta} \frac{P^i}{(1+r^i)^T}$$

$$v_t^i = \sum_{j=0}^{T-1} \frac{(1-\tau)y_{t+j,j}}{\prod_{l=0}^{j}(1+r_{t+l}^i)} + \frac{\bar{b}}{\prod_{l=0}^{T_I}(1+r_{t+l}^i)} + \sum_{j=T}^{\infty} \frac{(1-\delta)^{j-T}P_{t+j}^i}{\prod_{l=0}^{j}(1+r_{t+l}^i)} \tag{2}$$

Notice that the only source of individual risk in this economy is the agent's life span. Thus, the only reason for saving is to hedge the risk of outliving one's savings: there are only savings for retirement. We are abstracting from aggregate risk, which is not insurable in a closed economy, and other sources of idiosyncratic risk, like unemployment or health shocks during the working lifetime. From this point of view we are underestimating the amount of precautionary savings. Since all savings, independently of their original purpose, can be used in principle to hedge any kind of risk, before and after retirement, the bias would be small as long as the survival risk is sufficiently strong. As Gale and Scholz (1994) and Kotlikoff and Summers (1981) show, however, between 75% and 90% of individual savings can be explained by retirement reasons only. This assumption implies that we abstract from the risk premia embedded in interest rates, as our main focus is on the intermediation spread. When we calibrate the model in Section 4.1 we discuss how we adjust interest rates for risk premia to be consistent with this abstraction.

In order to capture in the simplest possible way the trade-off behind different strategies to hedge against retirement risk, we restrict the households' choice set to two alternatives: i) bank insurance or ii) self insurance. More importantly, we assume that households can only choose among these alternatives at age $j = 0$, and not at any other age $j > 0$. This assumption prevents all households from following the strategy of seeking high returns when young and switching to the strategy with better insurance just before retirement. This constraint has empirical support. Mankiw and Zeldes (1991) show ample evidence that most households do not ever hold stocks and prefer to keep all their financial assets in riskless alternatives (this is known as the participation puzzle) and even those households that hold stocks in their portfolios do not drastically change their strategies as they age.[10] Alternatively we could assume that changing the saving strategy upon retirement involves a cost, given that our setting is deterministic, this possibility would not change our results.[11]

To be concrete, households choose their retirement-saving strategy when they are born and, based on this decision, they choose the sequence of consumption both during their working life

---

[10]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.

[11]According to the IRS, individuals who withdraw from a retirement plan before 59.5 years old have to pay income tax on the amount taken out and an additional 10 percent tax.

and after retirement. The two retirement-saving strategies are:

1) **Strategy B: Bank-insurance:** Sign an annuity contract with a financial intermediary (a bank). An annuity contract between an agent and a financial intermediary specifies the payment that *the agent must make to the intermediary* during the agent's working age 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. The agent can choose to sign this annuity contract with two possible banks: a traditional bank (TB) or a shadow bank (SB). We will describe how these two types of banks differ in their liquidity costs in more detail later.

We assume that signing a contract with a shadow bank has an additional utility cost κ. This parameter captures several costs that are larger when via securitization, such as choosing and understanding these instruments, being subject to a higher probability of a crisis, etc. We model these costs in reduced form here just for expositionally and quantitative convenience, but in the Appendix we show how the results remain once the securitization effects on a higher crisis probability are taken into account

2) **Strategy S: Self-insurance:** Buy equity or bonds while working and live out savings after retirement, bequeathing any un-spent savings.

Remark on the Use of Annuity Contracts: We model financial intermediaries as only providing annuity contracts. This simplification is useful to introduce a clear insurance-return trade-off and to characterize sharply optimal choices. Annuities, however, should be interpreted as capturing investments that provide insurance more generally, including those that involve a wider set of assets and derivatives, and their combination. What is relevant for our purposes is that financial intermediaries offer either explicit insurance or a combination of assets that allows for its replication. The complexity of contracts and investment strategies that replicate the insurance provided by annuities is reviewed by Poterba (1997), who describes their provisions during both the accumulation and decumulation periods.[12]

---

[12]Examples include life annuities with payments over the lifetime, longevity annuities with a stream of payments that starts with a delay, a joint-and-survivor annuity with payments also to the survivor, and a "years certain" annuity, in which payments are guaranteed to continue for at least a certain number of time periods.

2.1.2 Productive Sector

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

$$Y_t = K_t^\theta (\Gamma_t L_t)^{1-\theta}$$

$$\Gamma_{t+1} = (1+\gamma)\Gamma_t$$

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.

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

$$y_t = F_L(K_t, \Gamma_t L_t)$$

where δ_k is the capital depreciation rate.

Life time-wage profile. Notice that Γ is labor-augmenting productivity. Thus, because average productivity grows at the rate γ per year, individual wages will also grow at the 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 (which is not valued by households), follows a committed debt policy D_t^G (which is independent of prices and quantities in the economy) and pays an average Social Security transfer of $\overline{P}_t$. The government collects taxes 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 individu-

als, $i \in \{B, S\}$. That is,

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

Balanced Growth Path. Since η, γ, τ and g are all constant, 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. We will also present the main results comparing changes across stationary equilibria. Nevertheless, in Section 4.3 we compute the transitions between equilibria.

Since in a stationary equilibrium $D^G_{t+1} = (1 + \hat{\gamma})D^G_t$, thus keeping revenue and spending constant, equation (3) implies that changes in the debt policy would have an impact on returns and aggregate quantities.

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, we simply denote the life pattern of consumption as c_j.

2.1.4 Financial Intermediation

The financial sector consists of perfectly competitive banks that offer annuity contracts to those households that follow strategy B to save for retirement. These banks specify the gross rate $1 + r$ that a household receives per unit of saving made during its working age. With these savings, the bank can invest either in "safe government bonds" that pay with certainty a unit gross rate $1 + r_L$ per unit of bond or in a continuum of "risky loans" that pay a unit gross rate $1 + \hat{r}_e > 1 + r_L$ per unit of loan, but only with probability $1 - s_b$, as with probability s_b the loan defaults and pays nothing. As the bank invests in a continuum of loans, a known fraction s_b of loans default and there is no ex-ante uncertainty on their return. Each bank takes the return of bonds (that is, r_L) and the risk-adjusted return on loans (that is, $r_e \equiv (1 - s_b)(1 + \hat{r}_e) - 1$) as given.

We denote the total financial intermediary's liabilities (deposits obtained from households) by D and all assets by A. We also denote the fraction of assets that the bank chooses to invest in loans by f. We assume that banks face liquidity considerations that put an upper bound on how much a bank can invest in loans while still obtaining deposits. More specifically, banks are subject to potential coordination problems, under which all savers (both agents that are

working and agents that are retired) can decide at any moment to withdraw their funds (a bank run). In such event, if a bank does not have enough funds to cover these withdrawals, it must default completely on all depositors.[13]

How easily can a bank liquidate its assets on short notice so as 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 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. There are many reasons why buyers cannot reap all the benefits of non-originated loan, which range from asymmetric information considerations under which the intermediary has superior knowledge about the quality of its own loans, to relationship lending that makes loans more easily monitored by the originator.

Then, for a given rate r promised to savers, the bank is *resilient* (not subject to a bank run) as long as

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

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

$$A = D.$$

Now we introduce the market for fire sales to describe the determinants of the liquidation

---

[13]This is naturally a very extreme assumption that can be relaxed and simplifies the exposition greatly.

rate q, in particular the role of securitization. We assume that, in case of distress, the bank randomly matches with another intermediary to sell its loans. Since the buyer may not have the expertise to operate the loans, it can try to rematch the loans and obtain the corresponding return r_e, with a probability

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

If the buyer does not find an intermediary that can operate the loan, then it does not obtain any return. This probability is assumed to be increasing in the amount of loans obtained (because of better pooling possibilities, for instance) and decreasing in the ratio $\frac{1+r_e}{1+r}$ (which is a measure of the specialization of the loan vis-a-vis government bonds or other standard assets). The probability is also increasing in an exogenous parameter $\Psi \geq 0$ that captures the technology available for finding counterparties and 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. Finally, the probability is also increasing in a parameter ζ that we just introduced to guarantee it 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 its main qualitative properties hold.

The demand for loans by a distressed intermediary 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 intermediary (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 $q_D = q_S$ and then $z^* = \frac{f}{1 + \Psi - f}$, subject to the constraint that

13

$z^* \leq f$, as the bank cannot sell more loans than it owns. As a result, the operation of this market puts a bound on the fraction of loans an intermediary can hold to guarantee enough funds for liquidation in case of a bank run. This constraint can be rewritten as

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

Each financial intermediary 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. 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). QED

Intuitively, when it is easy to trade assets (a liquid interbank market), there are fewer losses in case of liquidation and distress. The lower is the fire sale discount, 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 intermediaries can hold more productive assets in their portfolio and still successfully ride a run, zero profit conditions imply a better return for depositors (a higher f^* allows for a higher r^*).

Combining the equilibrium values for f^* and r^* we can define a risk-adjusted interest spread, ϕ, as

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

14

The risk-adjusted interest spread has two main components: 1) the physical cost of production, represented by the *value-added component*, $\widehat{\phi}$ and 2) the *liquidity component*. This last component depends on the securitization technology. It is zero when $\Psi \geq 1$ and increases as Ψ decreases (securitization becomes worse) otherwise.

Notice that in this model the liquidity constraint always holds but never binds, which implies that there is never a run in equilibrium and then the fire sale restricts outcome off-equilibrium. The absence of runs on the equilibrium path is an artifice borne out by abstracting from 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 traded easily in the interbank market so their value in case of liquidation is high (becoming a better substitute for bonds). We assume that some banks operate with Ψ_{TB} while others with $\Psi_{SB} > \Psi_{TB}$. We refer to the former as traditional banks and the latter as shadow banks. First, shadow banks have operated trading securities much more than traditional banks. Second, traditional banks face larger regulatory constraints, which put exogenous additional constraints on f. As is clear from the previous analysis, and essential for our quantitative exercise, shadow banks can invest a larger fraction of their portfolio in more productive loans, face fewer liquidity costs and offer a larger return to their investors.

2.1.5 Aggregates and Definition of (Stationary) Equilibrium

Here we define aggregate variables and the stationary equilibrium. We focus on steady state comparisons, but in Section 4.3 we study transitions and show that the comparison of steady states captures quantitatively most of the main benefits of shadow banking after the increase in life expectancy experienced in the U.S. since the eighties.

First, we specify aggregates along the balanced growth path. Distinguish by $i \in \{B, S\}$ agents according to their saving strategy. Since 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 $\mu_i = \int_{\alpha \in \mathcal{A}^i} m(\alpha) d\alpha$. As in every period t, a density $(1 + \eta)^t m(\alpha)$ of agents are born and their survival probabilities are exogenous; the density of agents of age j and type α who

choose strategy i is given by

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

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

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

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

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

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

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

where C is aggregate consumption along the balanced growth path; w^B and w^S are the individual net worths for agents following strategies B and S, respectively; W^B and W_t^S are the corresponding aggregates; B is the aggregate bequest; and L_t is total labor supply.

Definition 1 *Stationary Equilibrium.*

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

16

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

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

5. *Markets clear:*

- *Feasibility:* $Y = gY + C(r_e, \bar{b}) + X + \phi\left[\frac{W^B(r_e, \bar{b})}{1+r} - D^G\right]$ [14]

- *Assets market:* $\frac{W^B(r_e, \bar{b})}{1+r} + \frac{W^S(r_e, \bar{b})}{1+r^e} = D^G + K$

- *Bequest=inheritance:* $\bar{b} = (1+\gamma)^{T_I} B(r_e, \bar{b})$

2.2   Equilibrium Characterization

We solve the equilibrium backwards. First, we solve for the consumption path conditional on each saving choice for households of different α. Then we show their optimal saving decisions.

We first consider strategy B. The following analysis is regardless of whether the individual chooses to use traditional or shadow banking, as these cases will only change the received interest rate, r. Any household following strategy B would maximize the utility in equation (1) subject to equation (2). In the appendix we show that the solution is characterized by:

$$
\begin{aligned}
c_j^B &= \bar{C}^B \beta^j (1+r)^j v_0^B \\
b_j^B &= \alpha \bar{C}^B \beta^j (1+r)^j v_0^B
\end{aligned}
\tag{8}
$$

for some constant $\bar{C}^B > 0$. Notice that b can be considered as another consumption good, so that intra temporal optimality imposes $b = \alpha c$. Furthermore, households signing annuity contracts perfectly smooth consumption. For instance, if $\beta(1+r) = 1$ a household following strategy B would experience constant consumption throughout its life and would leave exactly the same bequest, independently of how long the household lives. This consumption plan implies the

---

[14]Note that in the feasibility constraint, what enters is the spread between the interest rates, ϕ, and not only $\hat{\phi}$. This is because f only makes financial intermediation more difficult, but does not stop it. Eventually, all savings must be intermediated; thus, if $f < 1$ the transaction takes more intermediation steps than when $f = 1$.

following pattern for the net worth of a household choosing strategy B:

$$
\begin{aligned}
w_0^B &= 0 &&\text{(9)}\\
w_j^B &= (w_{j-1}^B - c_{j-1}^B + (1-\tau)y_j)(1+r), && 1 \le j \le T, \ j \ne T_I\\
w_j^B &= (w_{j-1}^B - c_{j-1}^B + (1-\tau)y_j)(1+r) + \bar{b}, && j = T_I\\
w_j^B &= \sum_{t=0}^{\infty} \frac{(1-\delta)^{t-1}}{(1+r)^t}\left[(1-\delta)c_{j+t} + \delta\alpha b_{j+t} - ss_B y_T\right], && j > T
\end{aligned}
$$

Agents are born with zero wealth, and they work and deposit in the financial intermediary any non-consumed income, which generates a return r. At age T_I each household receives an inheritance, which is mostly saved; thus the net worth jumps at this age. After retirement, the financial intermediary pays the signed agreement and the net worth for the household is the present value of the contract.

Now we consider strategy S. Households in this case 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 when self-insuring solves

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

subject to

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

where r_e is the risk-adjusted return on equity.

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

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

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

$$c_j^S = \bar{C}^S \beta^j (1 + r_e)^j v_0^S, \qquad\qquad j < T \tag{11}$$
$$w_T^S = [1 - \sum_{j=0}^{T-1} \bar{C}^S \beta^j](1 + r_e)^T v_0^S$$

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

$$w_0^S = 0 \tag{12}$$
$$w_j^S = (w_{j-1}^S - c_{j-1}^S + (1 - \tau)y_j)(1 + r_e), \qquad 1 \le j \le T,\ j \ne T_I$$
$$w_j^S = (w_{j-1}^S - c_{j-1}^S + (1 - \tau)y_j)(1 + r_e) + \bar{b}, \qquad j = T_I$$

Two features of this economy are apparent when we compare equations (11) and (10) with equation (8). First, since $r_e > r$ before retirement, the consumption of self-insuring households grows faster that the consumption of bank-insuring households. After retirement, however, self-insuring households experience a faster decline in consumption than bank-insuring households. In fact, the consumption of self-insuring households converges to zero as the household lives long enough (see Figure 2). The difference in the return also has implications for the net worth distributions. Since the return on assets of self-insuring households is larger than the return of bank-insuring households, their net worth grows faster during the working life, but after retirement declines faster.

Figure 2: Lifetime Pattern of Consumption Under Strategies B and S
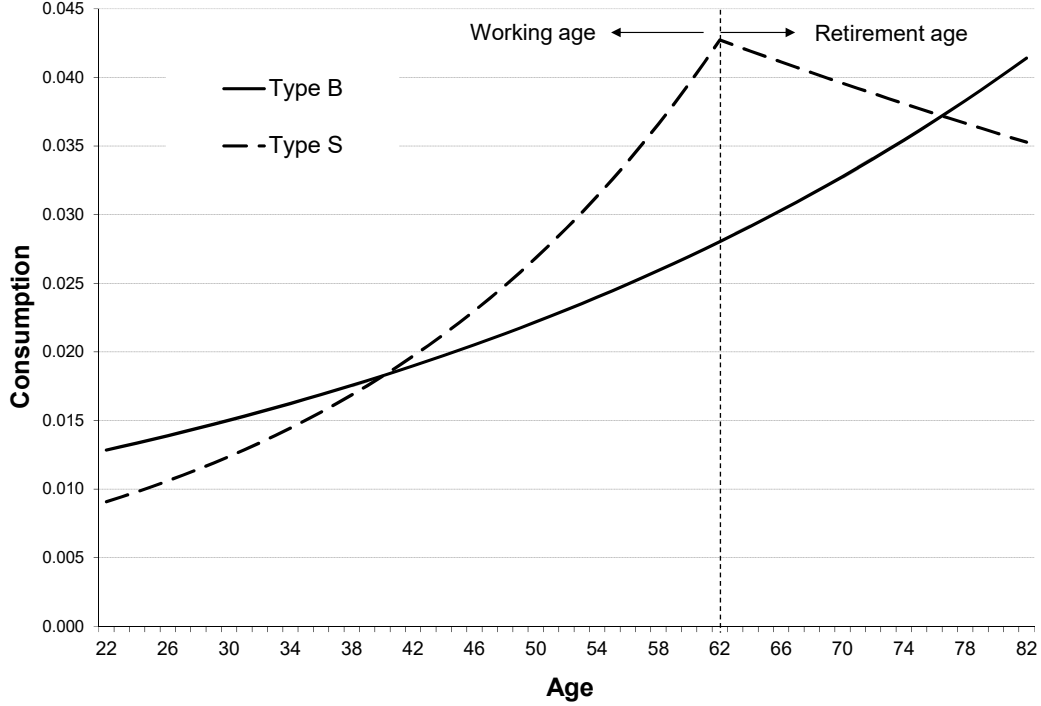
Now, based on these different consumption paths, we characterize the retirement-saving decision of households when entering the labor force. When a household chooses its retirement plan, it faces the following considerations. First, conditional on choosing strategy B, the household must choose whether to sign the annuity contract with a traditional bank or with a shadow bank. The trade-off that these two alternatives present is that the return from saving in shadow banks is higher, but represents a cost (of searching, understanding the contract or potentially facing a crisis, if we were allowing for aggregate risks). Since a utility cost κ is incurred at the time of signing the contract, the net present value of returns depends on the life expectancy (how many years the individual expects to have those returns). The next proposition shows that, conditional on signing an annuity contract, the agent chooses shadow banking as long as she expects to live long enough. This is true when the agents' bequest motive is not so large such that they would rather die faster in order to leave a bequest to their offspring. As we show later, however, the agents selecting into banking, and for which the decision is relevant, are those with low bequest motives.

Proposition 2. 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.

To clarify the forces driving the result in Proposition 2 we have assumed that κ is constant and independent of δ. As we mentioned before, this specification is a reduced form for either search costs, in which case the constancy of κ arises naturally, or additional risks involved in the SB activities. Under this last interpretation, one may wonder that κ could also depend on fundamental parameters. To address this issue in Appendix A.2 we consider an alternative environment where instead of the fix cost κ, upon retirement 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:

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

We can see that this cost is increasing in both the cost and the probability of the crisis (note that because $\zeta < 1$, then $\log(\zeta) < 0$). The above specifications also points out that the cost of SB increases with life expectancy, but as we show in Appendix A.2 the cost increases a smaller pace than the benefit. Thus, under some additional conditions (also satisfied in our quantitative exercise), we are able to prove an analogous result to Proposition 2.

Second, after determining which is the optimal annuity contract to sign given δ, households choose between strategies B and S. Strategy B has the benefit of fully insuring against the risk of living long, but it has the cost of generating a low return on assets. Conversely, strategy S has the benefit of generating a high return on assets, but it has the cost of not providing insurance against living too long. In particular, households following strategy S could leave large amounts of accidental bequests. Of course, the stronger is the household's bequest motive the lower the implicit cost of accidental bequests.

Proposition 3. 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 a full insurance technology, but some of them - those with large bequest motive- choose not to use it. They just self-insure. 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".

Using Proposition 2, from now on, and without lost of generality, we assume that the distribution of bequest motives is concentrated in two points: $\alpha = 0$ with probability μ and $\alpha = \hat{\alpha} > 0$ with the complementary probability $(1 - \mu)$. We will assume $\widehat{\alpha}$ is large enough such that these individuals do not change strategy when we perform quantitatively relevant changes in life expectancy, even though the individuals with $\alpha = 0$ who choose strategy B may change their preferred bank (traditional or shadow) with which to sign annuities.

Given the utility value of the available alternatives, each agent must decide his retirement strategy upon entrance into the labor market. It is clear that when $\alpha = 0$, the annuity strategy strictly dominates self-insuring, as $r_e \to r$. Thus, for $\alpha = 0$, there exists $\underline{\phi} = r_e - r > 0$ such that bank insurance is a better strategy. Further, as α increases, the value of both strategies increases. As long as $\frac{1+r_e}{1+r} \geq \beta \left[\frac{1-(1-\delta)\beta}{\delta\beta} \right]^{1-(1-\delta)\beta}$ the value of self-insuring increases faster than the value of bank insuring. As a result, as long as the interest differential is neither too small nor too large, there is a threshold for the bequest motive such that all households with a bequest motive below the threshold follow the annuity strategy, while the others self-insure.

The result of this assumption is that we will be able to change the composition of the banking industry between traditional and shadow banks but not the size of the banking industry. Endogenizing the size of the banking sector is beyond the scope of this paper.

3   Measuring Shadow Banking and Intermediation Costs

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

As there is no readily available measure of ϕ, as a proxy for intermediation costs we use spreads between interest received and interest paid in the financial sector from NIPA tables that encompass the whole financial sector. However, we have to make some adjustments. First, we have to acknowledge that productive investment opportunities are risky and some of those loans will not be recovered by the bank. For that reason we will adjust for "bad debt expenses" that subtract from the interest received. Second, when accounting for the rate paid by financial intermediaries to depositors and savers, we have to acknowledge that there are many other services provided that are not priced in, such as safety, accessibility to ATMs, financial advising, insurance, etc. For this reason we will adjust for "services furnished without payment," which adds to the interest paid.

To be more precise we want to measure $\phi = r_e - r$, where r_e has to be corrected from

defaulting debt and r has to account for non-priced services. As we discussed $r_e = (1 - s_b)(1 + \widehat{r}_e) - 1$, where $\widehat{r}_e$ is the rate charged for loans and s_b the fraction that defaults. We also define $r = r_L + r_s$, where r_L is the interest paid for savings and deposits (same as the price of bonds) and r_s is the return for other services not priced by banks. Then

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

The components of the last expression have counterparts in NIPA tables, which we measure as follows:

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

   This expression represents the average return on assets for all concepts that banks receive. To obtain this average 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.[15] To express these values as a return, Table D.3 of the Flow of Funds provides information for all the liabilities of the main economic sectors. Since we are interested in private borrowing and lending, we subtract the outstanding government debt, that is federal, state, and local liabilities. We call the resulting quantity of privately intermediated debt, just *hh's debt*.

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

   This expression represents the average return on deposits (or return on debt) that depositors and savers receive. Table 7.11, Line 4 of the NIPA tables provides information for the total interest paid on deposits by the financial sector, which we divide by privately intermediated debt as measured in the previous point.

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

   This expression represents 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 divide by privately intermediated debt as measured in the previous point.

---

[15]To account correctly for final intermediation, 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.

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 investments. The remaining fraction corresponds to *traditional banks* that only allocate a fraction $\widehat{f}$ of their assets to productive investments as they suffer the possibility of facing runs and coordination problems. The fraction of productive investments in the financial sector is then given by,

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

   As we discussed in the previous section, a part of the investments by traditional banks is also channeled by *shadow banking activities* (securitization, sponsoring of special purpose vehicles, participation in repo markets, etc.). In order to avoid double counting, 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. More specifically, we divide consumer credit from Table 110, Line 14 plus mortgages from Table 110, Line 15 by the total consumer credit and mortgages obtained by all households from Table D3 (columns 3 and 4). Then, we compute $\widehat{f}$ by the fraction of loans in the portfolio of traditional banks. More specifically, we divide all the loans from traditional banks (consumer credit, mortgages and others from Tables 110, Lines 12, 14 and 15) by all their deposits (checkable and time and savings) in traditional banks from Table 110, Lines 23 and 24.

Combining these components, Figure 3 shows the spreads since the seventies. As is clear, right before 1980 spreads were stable at around 4%. There was an increase in 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. Shadow banking institutions (s) increased from 5% in the seventies to more than 50% in recent years. Shadow banking activities also increased. These activities are captured in part by $\widehat{f}$, which rose from 80% in the seventies to almost 100% before the crisis, and then collapsed to 70% after the crisis.

Why did spreads decline? Have financial intermediaries increased their efficiency or improved their management of liquidity provision in the last decades? Philippon (2015) performs a thorough calculation of the changes in efficiency of the financial sector in the U.S. during the last 100 years using data on value added. He shows that $\widehat{\phi}$ has been constant for more than
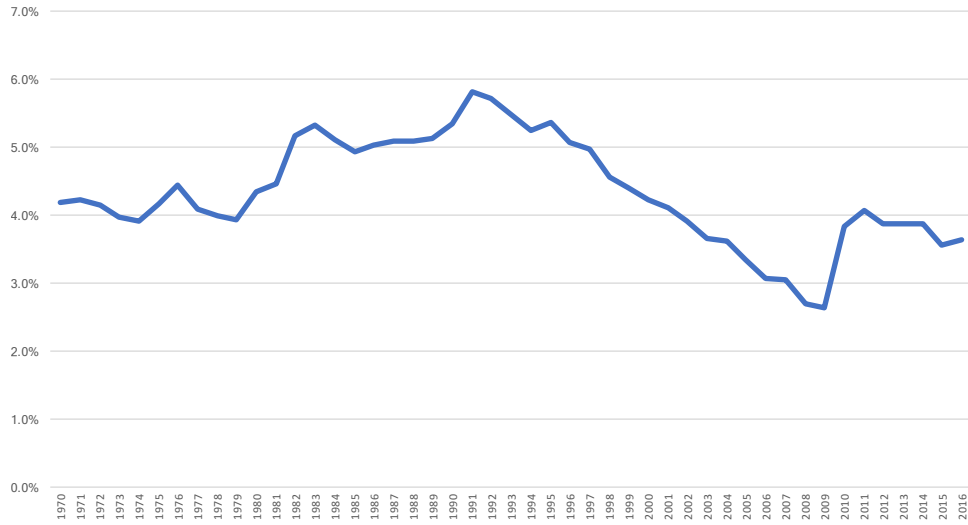
Figure 3: Risk-Adjusted Spread, ϕ

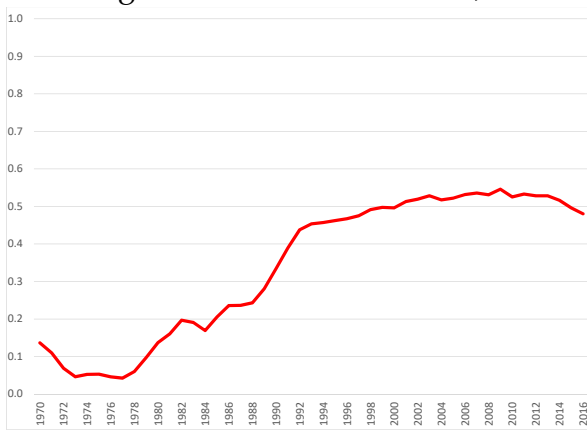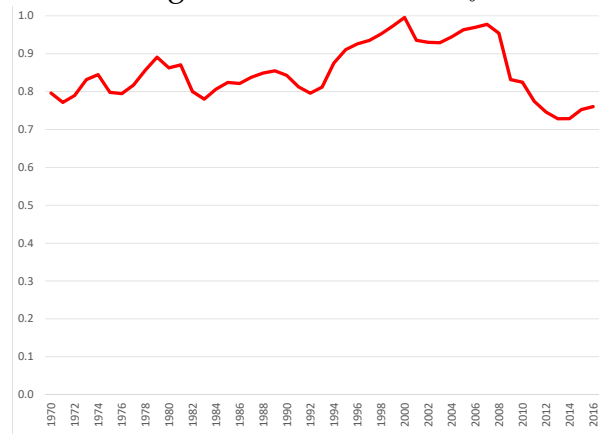## Figure 4: SB Intermediaries, s

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

100 years and that the technology in the financial intermediation industry exhibits constant returns to scale. He performs two alternative calculations: one assuming that the composition of the types of loans offered by the financial sector has remained stable during the sample period and another adjusting for changes in the quality of the loans. The first panel of Figure 6 shows the evolution of $\widehat{\phi}$ estimated by Philippon (2015). The takeaway from this picture is that the efficiency of the financial sector during the period under consideration has experienced little change.[16]

---

[16]When computing the per-unit value added, Philippon (2015) explicitly, and correctly, discards the use of the intermediation spread as a measure of value added. As we show in equation (7), the intermediation spread is affected by other factors that, even though they do not reflect physical costs, deeply affect the cost of financial

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

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

This is the difference between the realized spread and the potential one in that case where liquidity is not an issue and banks are able to invest all their assets in productive opportunities, that is, $f = 1$. The second panel of Figure 6 shows the evolution of the liquidity cost during the period 1970 to 2016. During the seventies the liquidity cost of intermediation was around 1%, and by 2007, the liquidity cost was almost 0%. After the recent financial crisis, the liquidity cost of intermediation increased again to almost 0.5%.

Figure 6: The Components of Spreads
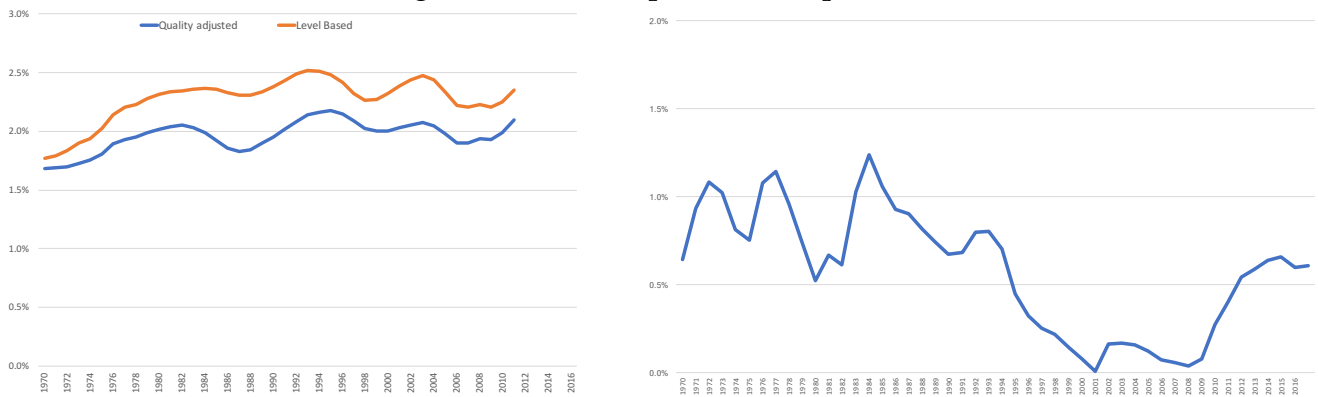
In short, the last three decades have been characterized by a large increase in borrowing and lending and by a large drop in the financial intermediation spread. All of the fall in the spread seems to have been led by an important reduction in the financial sector's "liquidity cost." Shadow banking has had a direct impact by making assets with higher returns useful as safe assets, replacing less profitable government bonds on banks' balance sheets and improving investment and output.

---

intermediation. As the focus of our paper is on understanding households' incentives to use the financial sector, the intermediation spread is, for us, of first-order relevance.

4 Quantitative Assessment of the Model

To perform a counterfactual experiment and decompose the macroeconomic effects of life expectancy (that led to an increase in the demand for safe assets) and the rise of shadow banking (that led to an increase in the supply of safe assets, which we claim was partly fueled by the increase in demand), we first calibrate the economy to replicate the main aggregates for financial intermediation in 1980. Then, we obtain the model's output for 2007, imposing newly observed life expectancy and intermediation costs. We 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 replicate yearly data. There are some parameters that are standard in the literature. These are (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, $\eta = 0.01$.[17]

When choosing the capital depreciation rate, δ_k, we need to take into account what capital means in our economy. In general, the literature targets a capital-output ratio of 2.7, which is the approximate ratio for the U.S. economy when one focuses only on productive capital. In our environment capital encompasses many other physical assets that constitute wealth for the household, such as housing and land. Including these assets the capital output ratio is about 3.4, which is generated by a depreciation rate of $\delta_k = 0.0271$.

Regarding the life cycle, we assume that agents enter the labor force at age 23. Since the average retirement age is 63.5 we set $T = 40$ (we only consider integer years) . In addition, using the Survey of Consumer Finances, we found that households receive inheritances on average at age 52. So, we set $T_I = 29$.

In terms of the parameters that determine fiscal policies, we obtain government spending as a fraction of GDP from the NIPA tables, $g = 0.20$. As for government debt, its 1980s ratio

---

[17]The calibrated β parameter is larger than standard values used by 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 has to 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 discount little 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.

(including state and local governments' debt) to GDP was around 0.40, while in 2007 the ratio was 0.62. This in principle implies a larger provision of assets that can be used for insurance purposes. A big part of this relative increase in government debt, however, was held by foreign investors. For our purposes, since it is the domestic availability of these assets what is relevant, we define the **net supply** of government debt as total government debt **minus** debt held by foreign investors. In 1980 the proportion of public debt held by non-U.S. residents was 20%, which implies a ratio $D^G/Y = 0.33$.

Given these choices there are three parameters left to calibrate: (i) the level of bequest motives of those households that care about their offspring, $\widehat{\alpha}$, (ii) the fraction of the last wage that the government transfers as Social Security after retirement, ss_i and (iii) the proportion of households that directly hold equity $1 - \mu$. Since there is no direct measure of these parameters, we normalize $ss_S = 0$ and choose ss_B, $\widehat{\alpha}$ and μ to replicate two important moments in the data: (i) the government debt to GDP ratio of 0.33 in 1980 and (ii) the private debt to GDP ratio of 1 in 1980. Since there are three parameters to match two moments, we choose the combination of parameters with the minimum μ (to be consistent with evidence from the Flow of Funds, as will become clear next). In this way we obtain $\mu = 0.75$, $\widehat{\alpha} = 4.64$ 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 Administrations.[18] Regarding the proportion of agents who follow each strategy, the Flow of Funds provides information about households portfolio choices, showing that 28% of American households directly hold equity, which implies that 72% hold equity indirectly, which is very close to our calibration with $\mu = 0.75$.

We will exploit two important parameters for the counterfactuals: (i) the survival probability after retirement, δ, which captures life expectancy and (ii) the spread between borrowing and lending, ϕ, which captures the role of shadow banking. 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, which is the observed value in 2007.[19] Based on Section 3, we calibrate $\phi = 0.04$ for 1980 and in

---

[18]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.

[19]These numbers can be computed using the actuarial tables of the Social Security Administration assuming that workers retire at an average age of 62.5. We use slightly conservative estimates, since life expectancy is still in

the counterfactual we decrease its value to $\phi = 0.03$, which is the observed value in 2007.

Finally, we can assess the model's performance for moments that we have not targeted. First, the model succeeds on the ratio of private consumption to GDP, as it generates a ratio of 0.56 in 1980, very close to the observed one of 0.62. Second, the model fails 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.

4.2   Decomposing Life Expectancy and Shadow Banking

We now show a counterfactual exercise for 2007 to decompose the effects of the change of life expectancy and the change in intermediation costs on asset accumulation, output and welfare, from 1980 to 2007.

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, as the ratio increased to 62%, but around 45% of the U.S. federal debt was held by non-U.S. residents.[20] Based on these figures we argue that the provision of government bonds during the period under consideration didn't play an important role in supplying safe public assets. Third, as in the data, we maintain the replacement ratio (that is, the proportion of wages obtained by 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 both the life expectancy increased (captured by a reduction in δ from 0.072 to 0.052) and the agents who signed annuities move from saving in traditional banks to saving in shadow banks. Because of Proposition 1, there exists two levels of search cost, $0 < \underline{\kappa} < \bar{\kappa}$ such that if $\kappa \in [\underline{\kappa}, \bar{\kappa}]$, it is optimal for those 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.

---

a downward path.

[20]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.

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' financial assets, a proxy for debt instruments. 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).

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 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 who save with annuities in terms of GDP (roughly constant at 1.3), while household debt over GDP would have slightly declined (from 1 to 0.96). Finally, the increase in retirement needs without an improvement in intermediation costs would have increased the demand for safe assets without an increase in supply, generating a reduction in their 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 toward shadow banking. This exercise shows what would have happened in our model with shadow banking but no extra needs for retirement since 1980. In this case, the increase in the capital to output ratio and steady state output would have been between 40% and 50% of the

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 S)	2.40	2.68	2.08	2.28
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/GDP	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 B	-	-	2.5%	2.8%
Plan S	-	-	-4.3%	-4.8%

total increase with the 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 for financial assets for higher life expectancy 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 funds without an increase in demand for retirement needs induces an

31

increase in the return on safe assets (r increases from 3% to 3.4%). Still, since more funds are channeled to investment opportunities, the equity return still declines (r_e declines from 7% to 6.4%).

Understanding the Decomposition: To build intuition about the sources of the previous decomposition, we show in Figure 7 the partial equilibrium 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 S agents (the net worth that agents of type S are willing to accumulate at a given interest rate r_e, given by $W^S(r_e)/(1+r_e)$,) depicted as the increasing solid curve without markers. As the return on capital r_e increases, agents of type S want to increase their savings, demanding more capital. The second component is the *indirect demand with borrowed funds*, which is the amount of capital that financial intermediaries demand on behalf of B agents. This second component, however, is determined by the operation of financial intermediaries and what 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 that agents of type B are willing to accumulate at a given interest rate r and cannot be held in government bonds (that is, $W^B(r)/(1+r) - D^G(r)$). These funds are deposited in financial institutions, which in turn lend it to equity holders. This is what we called before the indirect demand for capital. The solid decreasing red curve, $\Upsilon(r+\phi) \equiv K(r+\phi) - W^S(r+\phi)/(1+r+\phi)$, is the demand for credit to buy capital, which is the capital that cannot be bought by agents of type S with their own funds. As is clear, both markets are interlinked and cannot be solved separately. In this expositional abstraction, when we explain one market, we take the other as given.

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 (i.e, δ changes from $\delta_H = 0.072$ to $\delta_L = 0.052$)? This counterfactual is shown with dotted lines. First,

32

Figure 7: Partial effects
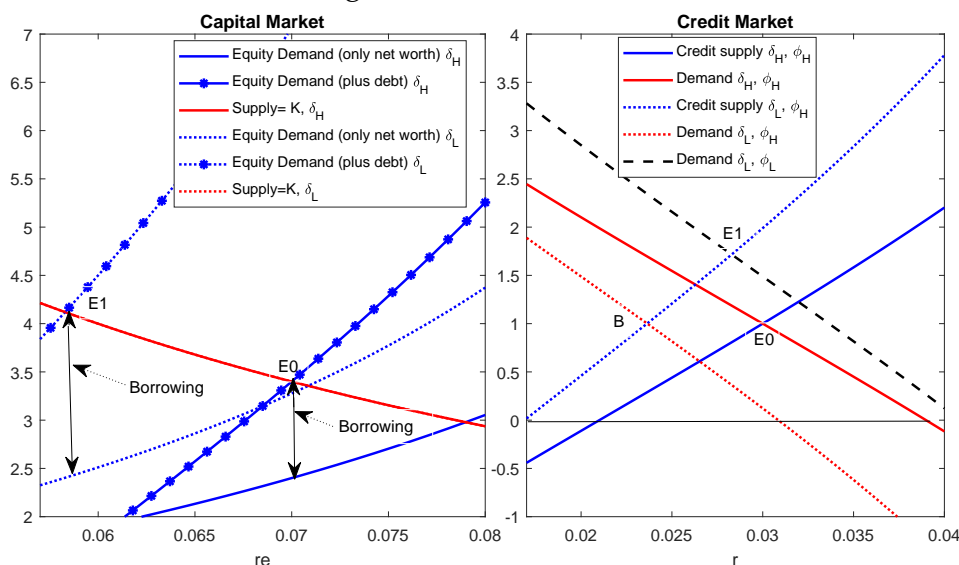
we discuss what happens in credit markets. Since B-agents expect to live longer, they accumulate more assets that must eventually be intermediated. This is represented by an increase in the credit supply in the credit markets. There is also a decrease in the demand for credit by S-agents. As their life expectancy also increases, they accumulate more assets on their own and then require less 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 ($r = 0.022$). What happens in capital markets? While the supply of capital is not affected (it is purely a technological function), the demand increases for two reasons: S-agents expect to live longer and they save more to buy capital.

The previous discussion shows that the rising demands for insurance of 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 baking in this picture tends to have a first-order effect in 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 implies that more resources can be used per unit of investment opportunity (as $\Upsilon(r + \phi_L) > \Upsilon(r + \phi_H)$, for each r) and there is an increase in the demand for credit by S-agents. With the introduction of shadow banking, which makes the functioning of credit markets more efficient, the new "partial" equilibrium is at point $E1$, with a higher intermediate interest rate r, a higher level of credit and an even higher level

of capital, all of which could not be generated by just changing δ.

We keep emphasizing the partial nature of the previous 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 that are imbedded in the construction of these curves. 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.

Remarks on Welfare Effects: When there are changes to "preferences" (in our case life expectancy) affecting the computation of present values, comparisons across experiments become hard to interpret in terms of welfare. This is the reason we make comparisons fixing δ and using consumption equivalent changes. With logarithmic utility, as we have assumed, the calculations are quite simple.

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 $\widetilde{\underline{C}} = \{\tilde{c}_t, \tilde{b}_t\}_{t=0}^{\infty}$ 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:

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

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

From now on, the change in welfare is expressed in terms of $100\lambda\%$.

Comparing columns 1 and 3, 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%, S-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. In this case welfare increases by 0.4%. Again, the big gain in welfare comes from B-agents, who benefit from the more efficient financial system in the economy by 2.8%, while S-agents become worse off, experiencing a loss of 4.3% in terms of lifetime consumption.

Remarks on Flexible Government Debt/GDP: In the previous simulations we maintained D^G fixed. In Table 2 we consider alternative scenarios for 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 the life expectancy had increased, the spread had decreased to 3% and the government were allowed to freely choose the level of debt without changing either taxes or transfers. 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 experiment 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	1.06
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 S)	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/GDP	1.00	1.62	1.63	1.49
Data (FF: Table D3)	*1.00*	*1.66*		

By changing the 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, Primiceri, and Tambalotti (2013) and Justiniano, Primiceri, and Tambalotti (2015) relate the credit boom experienced during the 2000s to the international savings glut, and claim that the fall in interest rates was due to the influx of foreign funds. In particular, Justiniano, Primiceri, and Tambalotti (2013) argue that the international savings glut can account between one-fourth and one-third of the increase in U.S. household debt. The last column assumes that the debt to GDP ratio moves from 0.33 (as in 1980) to 0.62, which would be the domestic supply of government bonds in 2007 if foreign nations did not hold any U.S. Treasuries (see Section 4.1). This allows us to decompose the effect of the international savings glut proposed by the literature and the domestic savings glut that we introduce in this paper.

The direct effect of more government debt is an increase in the supply of safe public assets that maintain r at levels closer to the case without an increase in life expectancy, even though intermediation costs decline. This induces a decline in the 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 U.S. household debt. 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 output and in the capital-output ratio 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 a life expectancy change in the economy from 1980 to 2007. In this section we show that the convergence indeed happened quite fast: by 2010 most of the increase in debt (around 90%) had already happened. We also show that cyclical movements of economic activity play an important role in accommodating the slow growth in debt during the early 80s and the subsequent speeding up in private indebtedness during the 2000s.

Figure 8: Transition Dynamics: Constant TFP

(a) Interest rate

(b) Output

(c) Assets

(d) Household Debt/Output
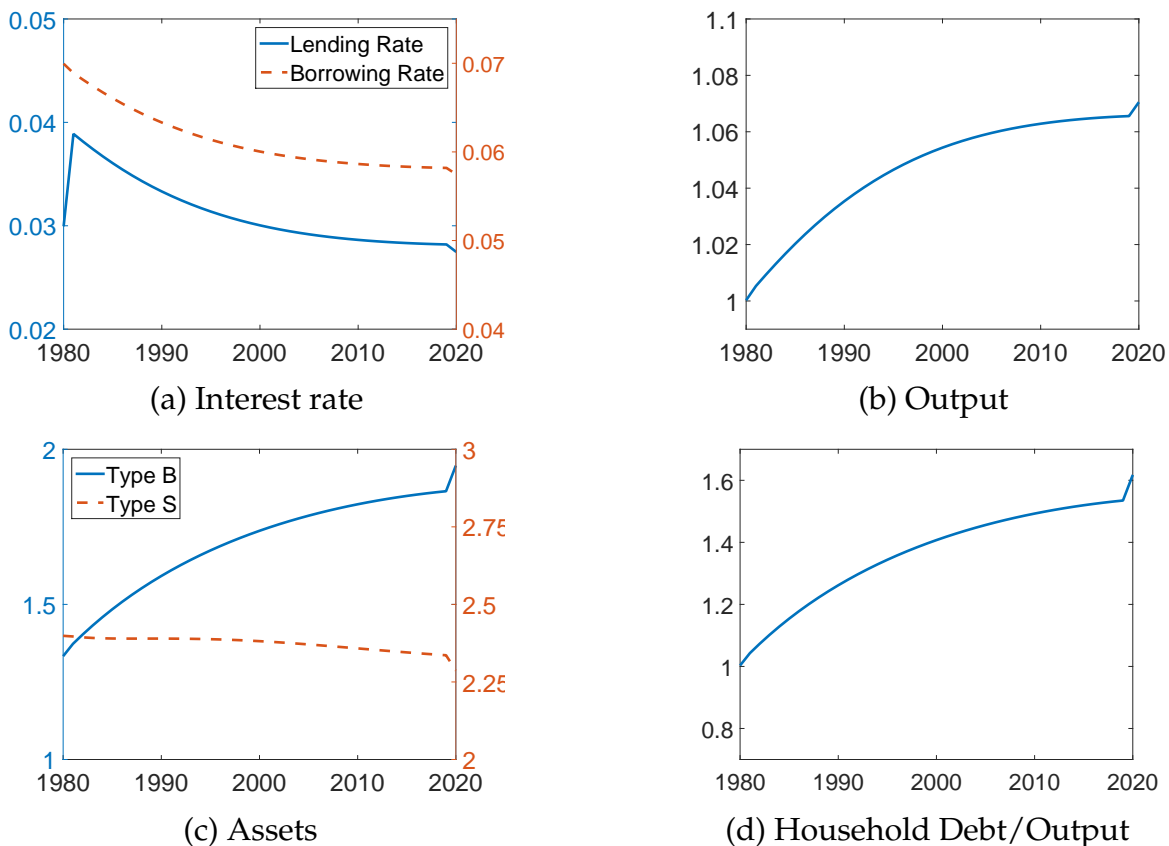
The computation of the transition, however, presents several challenges. First, at the time of the shock, there is a distribution of agents indexed by age and assets. Who is affected when the life expectancy shock happens in 1980? In what follows, we assume that agents who were retired at the moment of the shock continue to have the initial survival probability $\delta = 0.072$, while the survival probability for all working-age agents jumps to $\delta = 0.052$. Second, some agents were already in an annuity contract. What happens with those contracts? We assume that after the shock all of the existing contracts are renegotiated to take into account the new survival probability. Third, what happens with the government budget? We assume that the 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 the government's budget balance. Finally, what happens with retirement payments? We assume that for those already retired in 1980, these payments remain at the same level at which they were contracted before the shock.

In Figure 8 we show the transition dynamics of a shock that reduces δ from 0.072 to 0.052 and reduces ϕ from 0.04 to 0.03. In panel (a) of the figure, we see that the spread (the difference between lending and borrowing rates) falls drastically on impact and then increases 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 increasing net worth of B-agents. Interestingly, the net worth of S-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 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
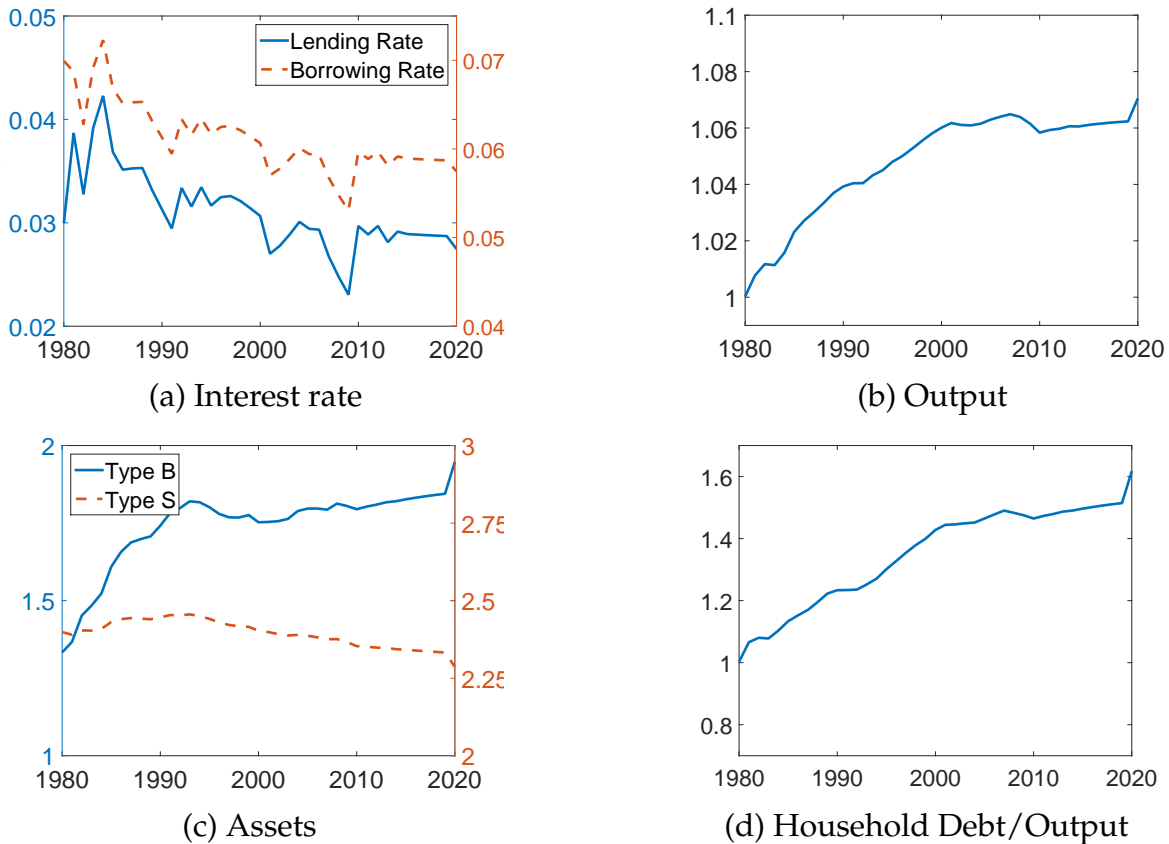
In Figure 9 we compute the transition path using the actual path for TFP (measured by the Solow residual) instead of fixing $A = 1$ as in the previous simulation. This exercise is informative because it shows how the recessions in the early 80s and early 90's slow down the convergence during the 80s, and how debt speeds up in the second half of the 90s. Although the evolution of interest rates is mostly unaffected by the changes in TFP, output and private debt are slightly above the figures computed in Figure 8.

Remark on the Growth of Shadow Banking: Notice that, given our counterfactuals just rely on whether shadow banking exists or not, we are not required to calibrate κ because those corner situations imply a region for κ. It would be possibly, however, but 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 towards shadow banking along the transition as observed in the data. The speed at which shadow banking is adapted would then discipline such distribution of κ.

5   On the Costs and Benefits of Shadow Banking

This paper's goal was understanding the rise and benefits of shadow banking in the economy, thus abstracting from its potential cost in the case of inducing a crisis. The great recession, characterized by a collapse in securitization and other instruments used to provide liquidity at a lower cost in shadow banking activities, is a reminder of how large these costs. Using a methodology proposed by Luttrell, Atkinson, and Rosenblum (2013) and later expanded by Ball (2014) and Fernald (2014) that compares the potential output computed by the Congressional Budget Office (CBO) and the realized output, we computed 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, which is realized output after 2007.[21]

If shadow banking was solely responsible for 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 equivalent

---

[21]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 revised down potential output in the last two years.

Figure 10: The Costs and Benefits of Shadow Banking
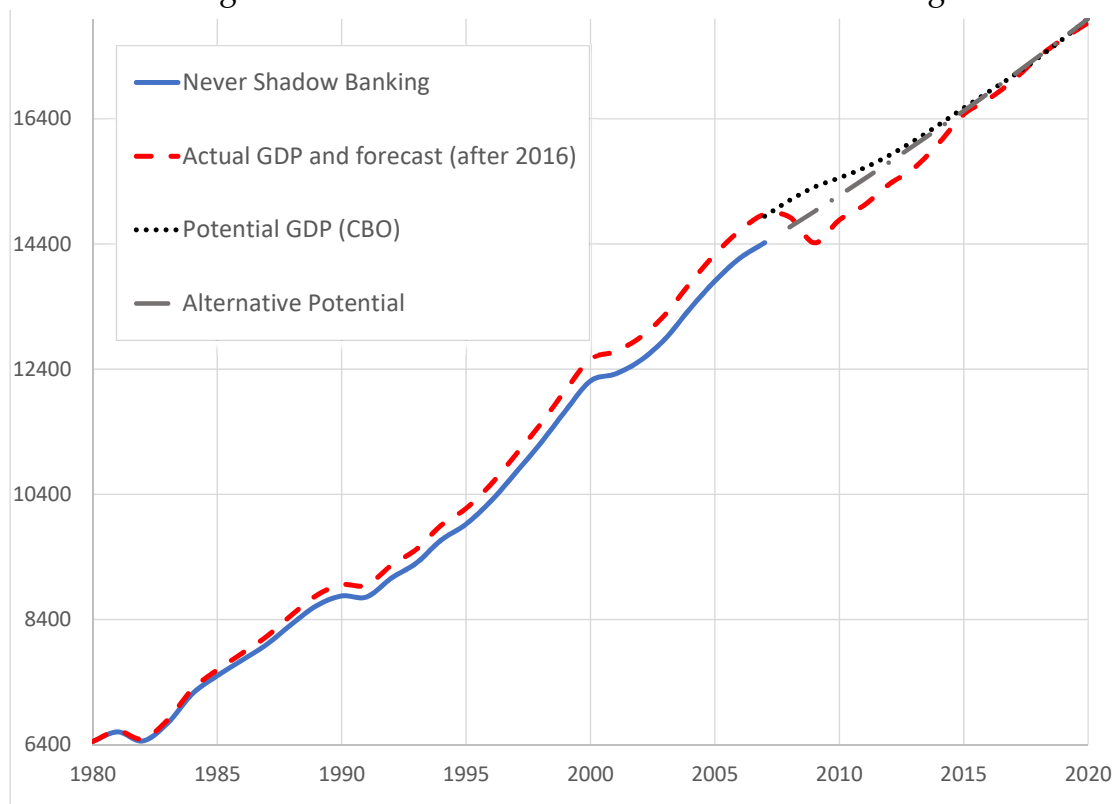
to 2.8% per year in the stationary equilibrium ($1.062 - 1.034$, according to the second and fourth columns of Table 1), which represents a present value of around 3.3GDP of 2007. This is, however, misleading. First, it assumes shadow banking is permanent and does not generate crises. Second, it overestimates 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, that is, without the potential gains from lower spreads but also without the cost of a crisis. For this counterfactual we assume $\kappa > \bar{\kappa}$, so that after the increase in life expectancy, individuals just keep choosing traditional banks and the spreads remain at the 1980 level of 4%.[22] 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

---

[22]In this counterfactual there is no shadow banking previous to 2007, and as we assume no crisis in the absence of shadow banking, we assume that the economy maintains the average growth level since 2007.

output from 1980 to 2007 (which, as we discussed, corresponds to a transition period) 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 for 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 that 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 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.

We show that a calibrated model with an increase in the demand for safe assets for retirement needs and shadow banking that reduces the cost of financial intermediation of the magnitude observed in the data accommodates well the large increase in asset accumulation and output experienced by the United States since 1980. We find that in the absence of shadow banking assets would not have increased, while the capital-output ratio and output would have increased only half of what they did.

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 of 59%

of 2007 GDP. Further, even if we assume shadow banking was solely responsible for the great recession, its cost is on the order of 13.5% of 2007 GDP. In other words, 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 discussion about whether and how to regulate the banking system. Although avoiding 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 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 large regulatory changes.

References

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

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.

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

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: Economic Consequences of Demographic Change in a Global Economy April 21-22, 2006.

Bloom, David E., David Canning, and Michael Moore. 2014. "Optimal Retirement with Increasing Longevity." *Scandinavian Journal of Economics* 116 (3): 838–858 (07).

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.

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

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.

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).

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.

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

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.

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.

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 non-stockholders." *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.

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.

Poterba, James. 1997. "The History of Annuities in the United States." NBER Working Paper 6001.

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

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

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:

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

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

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:

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

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 $\delta > 0$.

A.1   Lemma 1:

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

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

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

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

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

47

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

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.

A.2 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 **?**, 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 $\underbrace{\alpha\left(1 + \log\left(\frac{1+r_{SB}}{1+r_{TB}}\right)^{T-1}\right)}_{>1} < \frac{1}{1-\beta} - \underbrace{\frac{\sum_{j=0}^{T-1}\beta^j \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]^j}{\beta^T \log\left[\frac{1+r_{SB}}{1+r_{TB}}\right]}}_{>0}$.

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.

B   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 $U_S(\alpha)$ represent the maximum attainable utility of an agent of measure zero in this economy who follows strategy B (baking) or S (self-insurance) respectively as a function of α. Define $\Delta(\alpha) = U_S(\alpha) - U_B(\alpha)$. 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 $\alpha \geq 0$.

B.1   Lemma 2:

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

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 follows self insurance strategy (S) is

$$U_S(\alpha) = \sum_{j=0}^{T-1} \beta^j \log(c_j^S) + \beta^T [\phi_S(\alpha) + \theta_S(\alpha) \log(w_T^S)]$$

where

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

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

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

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

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

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

Because the first term is independent of α it follows that

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

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

50

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

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

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

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

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

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

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

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

51

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 $\frac{\partial\Delta(\alpha)}{\partial\alpha} > 0, \forall\alpha \geq 0.$ ∎

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