



## Research Article

Mauro Caselli\* and Babak Somekh

# Access to Banking and the Role of Inequality and the Financial Crisis

<https://doi.org/10.1515/bejeap-2020-0421>

Received December 7, 2020; accepted September 23, 2021

**Abstract:** We study access to banking and how it is related to banks' rate of return on investments and the distribution of income. We develop our empirical framework through a theoretical supply-side model of bank deposit services with a consumer population heterogeneous in income. We use this model to show how decreases in the interest rate margin and higher income disparities lead to an increase in the proportion of unbanked. Using localized US household data from 2009, 2011, 2013 and 2015 we find strong empirical evidence for the predictions of the model. We then structurally estimate our model to estimate the value of having a checking account relative to alternative financial services and to quantify the effects of actual changes in the interest rate margin and the distribution of income that occurred in the aftermath of the 2008 financial crisis.

**Keywords:** financial exclusion, income inequality, financial crisis, alternative financial services

**JEL Classification:** D11, D12, D14, D21, D22, G21

## 1 Introduction

This paper studies financial exclusion and its determinants by taking advantage of a new dataset focused on the unbanked and underbanked in the US (FDIC 2017).

---

\***Corresponding author: Mauro Caselli**, School of International Studies and Department of Economics and Management, University of Trento, Via Tommaso Gar 14, Trento, TN 38122, Italy, E-mail: [mauro.caselli@unitn.it](mailto:mauro.caselli@unitn.it). <https://orcid.org/0000-0002-1132-1665>

**Babak Somekh**, Baruch College, CUNY, New York City, USA

In particular, we are interested in how access to banking is related to banks' rate of return on investments and the distribution of income of the consumers in the banks' target market. Our results show that when a bank faces a low interest environment it is forced to raise fees, pushing low-balance consumers out of the market. The impact of these fees on the rate of unbanked is exacerbated by the level of inequality in the target market. These results persist even after controlling for other market forces such as bank location and competitiveness of the market for deposits.

Understanding the determinants of financial exclusion is of paramount importance given that access to banking services is increasingly a prerequisite for participation in the mainstream economy (Jackson 2019). Households rely on bank accounts to conduct basic financial transactions, build precautionary savings, and as a means to access affordable credit. Most workers in advanced economies are no longer paid in cash, and require a way to cash checks or set up direct deposits in order to "access" their income. Almost all social assistance and unemployment benefits require a bank account, an issue of paramount importance to low-income households, as the COVID-19 pandemic made clear.

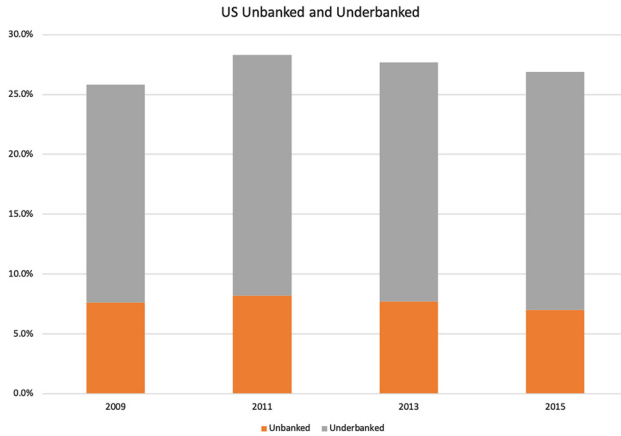
However, according to the Federal Deposit Insurance Corporation (FDIC), over one quarter of households across the US are either unbanked (6.5%, do not have a checking account) or underbanked (18.7%, have a checking account, but use alternative financial services such as prepaid debit cards and check cashing services) (FDIC 2017). Lack of access to banking is especially stark amongst low-income households and amongst minority groups in the US.<sup>1,2</sup>

Consumers that do not have a bank account usually turn to alternative financial services (AFS) for their banking needs. These services tend to be more loosely regulated and charge significantly higher fees than mainstream banks. This would suggest that those without a bank account may end up paying more for basic financial services, they may be more vulnerable to loss or theft of their cash and

---

**1** Financial exclusion is also a problem of varying degrees across Europe, with the rate of unbanked estimated at 1% in France and Germany and up to 10–15% in Portugal and Italy (Demircuc-Kunt et al. 2015; Parekh, MacInnes, and Kenway 2010; Rowlingson and McKay 2017). These figures are high considering that a recent directive from the European Parliament has established access to a bank account as a right, which has forced banks to provide low-cost no-frills accounts with limited fees (Rowlingson and McKay 2017). The issue is even more pronounced in poorer developing countries (Beck, Demircuc-Kunt, and Peria 2007, 2008).

**2** Looking at the UK, the rate of unbanked is estimated to be around 3%, but it used to be much higher (FSA 2000; Kempson and Whyley 1998). Paradoxically as the percentage of unbanked has decreased the AFS industry in the UK has seen a high level of growth.



**Figure 1:** Unbanked and underbanked in the US.

asset holdings and often have difficulty building credit histories and achieving financial security.<sup>3,4</sup>

Despite an increase in interest in financial exclusion from policy makers, identifying its determinants has been difficult due to lack of granular data on the unbanked across time for any country. Reliable data on financial inclusion for the US population have only become available in 2009 following a biennial program conducted by the FDIC (FDIC 2017). Figure 1 presents how the rates of unbanked and underbanked have moved over the last four iterations of the FDIC data. In our analysis, we propose to analyze why there may be links across the rate of unbanked, banks' interest margin, i.e. the difference between the rate a bank earns on deposits and the rate it pays to consumers for those deposits, and the dispersion of income, and to use the FDIC data to test these proposed links.

In particular, we set up our investigation by analyzing a theoretical supply-side model of bank deposit services with a consumer population heterogeneous in income. We construct our model based on the product differentiation framework of Shaked and Sutton (1982) and a model of exclusion presented by Atkinson (1995).

<sup>3</sup> Financial exclusion is also a public policy concern. For example, social security, unemployment benefits and other benefits payments made by government institutions usually come in the form of electronic payments or checks. To the extent that those receiving these benefits have to pay high access fees to AFS providers, this is a transfer of public assets to these financial institutions.

<sup>4</sup> Exclusion may also have a psychological cost, impacting consumer behavior (see, for example, Carrell and Zinman 2014).

This allows us to focus on the type of consumers that are excluded. We use this model to identify the mechanisms through which the interest margin and income dispersion impact access to retail banking and to develop testable predictions.<sup>5</sup> Our theoretical model also provides us with the specific structure we use in our maximum likelihood estimation described below.

Using localized US household data from 2009, 2011, 2013 and 2015, we find empirical evidence for the predictions of the model and in favor of our focus on the interest rate margin and the distribution of income. In particular, the results from the reduced-form analysis show that decreases in the interest rate margin and increases in income disparities lead to increases in the proportion of unbanked and higher financial exclusion.

Next we structurally estimate the model using maximum likelihood. This allows us to estimate the value of having a checking account relative to alternative financial services. Banks are estimated to provide a small but significantly positive benefit to low-income account holders relative to the use of AFS (as well as relying purely on cash transactions). This is not an obvious result. It may be argued that AFS providers are more convenient for low-income households, because of their accessibility and pay-per-use pricing (FDIC 2017). Our results suggest that there is a cost to being excluded. Moreover, our structural estimation makes it possible to quantify the effects of the actual changes in the interest rate margin and the distribution of income, occurred in the aftermath of the 2008 financial crisis, on financial exclusion. We argue that the recent financial crisis, through its impact on the interest rate, may have led to a significant increase in the percentage of households without a bank account.

This paper is closely related to the literature studying the reasons for financial exclusion. In particular, it is related to those few papers that show how banks play a role in determining the level of exclusion through their accessibility and pricing structure. Celerier and Matray (2019) examine how competition affects unbanked households. They show that increased competition after deregulation in the US led to higher branch density and caused previously unbanked households to open new bank accounts, especially in historically excluded areas. On the

---

<sup>5</sup> Banks charge for their services directly through fees and indirectly through foregone returns. While indirect fees are more costly for high-balance customers, high direct fees are a regressive pricing mechanism and are a way for banks to avoid less profitable customers (Bord 2017; Dash 2011; Son and Tighe 2011). In the global recession spurred by the financial crisis of 2008, banks were faced with declining returns on customer deposits as well as greater financial scrutiny and regulation on their investment portfolio. This led most US retail banks to target more profitable customers by raising the fees charged on low-balance accounts. This suggests that there could be a link between direct fees, indirect fees and access to banking.

other hand, Bord (2017) studies the impact of a related supply-side mechanism, i.e. consolidation and the emergence of large banks, irrespective of market concentration. They show that consolidation, which also partially resulted from the US deregulation laws, led to the predominance of large banks with higher fees and an increase in the unbanked households. Our paper looks to better understand the supply-side causes of financial exclusion by focusing on the profitability of low-income customers for banks and how such profitability is affected by the interest rate margin and the distribution of income. In particular, we add to the existing theoretical literature by explicitly considering the role of income distribution in the pricing decision of banks. We also include the aspects analysed in Bord (2017) and Celerier and Matray (2019) in our reduced-form analysis and we find that competition, measured by the number of branches, and concentration, measured by the share of deposits of the top five banks, do not affect the proportion of unbanked after controlling for the interest rate margin and the distribution of income.<sup>6</sup>

We also contribute to the empirical literature estimating the benefit of being banked. Dick (2008) and Ho and Ishii (2011) estimate the effect of location and competitiveness (due to deregulation) on consumer demand and welfare using the Berry, Levinsohn, and Pakes (1995) framework. Our theoretical model provides an alternative framework to structurally estimate the benefit of being banked relative to relying on AFS. We use the recent FDIC data to implement this framework and find a small but significantly positive benefit to having a standard checking account.

In the following section, we present a supply-side model of the market for banking services and develop several testable predictions. In Section 3 we describe the data we use in our empirical analysis. Section 4 provides reduced-form estimations of the main predictions of the model. In Section 5 we present a structural econometric model based on our theoretical framework and we show that the model fits the data well. We present counterfactual analyses in Section 6.

## 2 Theoretical Model

Our model considers how the interest rate margin and the distribution of income of consumers can impact bank decisions, focusing specifically on the supply of deposits in the banking sector. By introducing consumers heterogeneous in

---

<sup>6</sup> Others have looked at the role of information on the level of unbanked (see, for example, Bertrand and Morse 2011, and Stango and Zinman 2014), with inconclusive results. We do not assume any information asymmetry in our theoretical model.

income we are able to focus on some of the key causes and extent of financial exclusion in bank deposit services. We consider how banks price for deposit services and so how they determine the type of consumers that they accept deposits from. In our setup, the existence of alternative financial services in the market, i.e. financial services provided outside traditional banking institutions,<sup>7</sup> provides consumers with a better outside option relative to relying solely on cash for their day to day existence. In that sense the AFS market plays a positive role in our model, forcing banks to price more competitively.

We abstract away from the monitoring problem, taking the return banks earn on deposits as given, and focus on the cost-benefit tradeoff of the banks and deposit customers. The general framework of our model and our method of telling the story of the bank deposit market follows that of Shaked and Sutton (1982), who consider entry and the choice of quality in a monopolistically competitive market, and Atkinson (1995), who considers the exclusion of consumers from the market of a productive good. We have adjusted their assumptions about consumer preferences and firm strategy to reflect more closely the market for financial services.

## 2.1 Consumers

There is a unit mass of consumers that only differ in their income,  $w$ . Income is distributed according to a cumulative distribution function  $G(w)$ . The density,  $g(w)$ , is zero for values of  $w$  below the minimum income,  $a$ , and above the maximum income,  $a + h$ , where  $h$  can take any positive value,  $h > 0$ . Consumers can choose to either keep their earnings at a mainstream bank providing all the deposit services described above, or to turn to an AFS that offers a minimum set of services (such as check cashing or pre-paid debit cards). More formally, banks provide consumers with full access to their earnings as well as an additional benefit of  $\theta w$ , to a customer earning  $w$ .  $\theta > 0$  captures the positive value of having a bank account relative to relying on AFS. Banks charge a fee,  $f_B$ , for these deposit services. AFS only provide consumers with access to their earnings (this is analogous to  $\theta_A = 0$ ) and charge fee  $f_A$ , which corresponds to the total cost paid by its customers.

We are implicitly assuming that banks charge an indirect fee by not providing a deposit interest rate to customers, but we are not including this type of fees in the consumers' problem. This is based on the observation that most consumers

---

<sup>7</sup> Examples of alternative financial services include check cashing, pre-paid cards, money orders and remittances.

that use AFS providers do not have access to a risk free rate,  $r_f > 0$ , as an outside option. In addition, over the timeline of our analysis most banks have offered a zero interest rate on checking accounts. Therefore, we believe that these fees are not a real consideration when choosing between using AFS and maintaining a transaction account with a bank.

We assume that consumers do not have full access to their cash without going through a financial service provider. Otherwise AFS customers would choose to keep their income  $w$  and not pay a fee. This is based on the observation that in the modern economy most workers are paid through checks or direct deposits. Social assistance and unemployment benefits are also paid as checks or electronically. In addition, many consumer transactions, from online purchases to sending money to family members, usually require bank/AFS services.<sup>8</sup>

The consumer's binary choice is between:

$$u_B = (1 + \theta)w - f_B \text{ and } u_A = w - f_A \quad (1)$$

We compare the two utility functions above to determine the income level,  $w^*$ , such that consumers earning an income below  $w^*$  choose to use an AFS over a mainstream bank:

$$w^* = \frac{f_B - f_A}{\theta} \quad (2)$$

Consumers earning below  $w^*$  are considered excluded, or priced out, from mainstream banking services. We are particularly interested in looking at how the proportion of consumers that are excluded,  $G(w^*)$ , is determined within our model.

## 2.2 Banks

We use a version of the Monti-Klein model of a monopolistic bank, focusing on the deposit side of the bank's problem. We lose no generality in assuming that the bank is a monopoly, all of our results below would hold in a model of imperfect competition between a finite number of banks as described in Freixas (2008).<sup>9</sup> We work with the monopolist version purely for expositional reasons. The setup of

---

<sup>8</sup> This assumption is made in order to simplify the model and to focus the narrative. It is possible that some consumers at the bottom of the distribution completely rely on cash and never use AFS or banks. The existence of such consumers would not have an impact on any of our theoretical results presented below.

<sup>9</sup> Freixas (2008) show that the Monti-Klein model can be viewed as a model of monopolistic competition with limiting cases of monopoly and perfect competition. In our model the (finite) number of banks that share the deposit customers would only have a multiplicative impact on the first-order condition, and so we can focus on the monopoly case.

our model is not too far away from the duopoly setup considered in Gabszewicz and Thisse (1979) and the monopolistically competitive model of Shaked and Sutton (1982).

The bank takes in deposits and uses those deposits to invest in projects earning an assumed rate of return,  $r$ . The bank faces a fixed cost per deposit account associated with the administration and servicing of these accounts,  $c_B$ . Thus, the profit function for the bank is  $\pi_B = rD_B + f_B N_B - c_B N_B$ . Substituting for  $f_B$  from Eq. (2), the profit function for the bank becomes:

$$\pi_B = rD_B - (c_B - \theta w^* - f_A)N_B \quad (3)$$

$$\text{where } D_B = \int_{w^*}^{a+h} w g(w) dw \text{ and } N_B = 1 - G(w^*)$$

$D_B$  is the total amount of deposits taken in by the bank and  $N_B$  is the number of bank accounts.<sup>10,11</sup>

The demand faced by the bank is determined by the level of income where consumers are indifferent between the bank and an AFS provider,  $w^*$ , as defined in (2). In most general equilibrium models of bank deposits, such as Basu and Wang (2007), the rate of return is endogenously determined. These papers look to see the role of monitoring on the bank's demand for deposits. We are not interested in monitoring and so we take  $r$  as an exogenous parameter in our model indicating the economic conditions faced by banks and consumers.

We assume no barriers to entry in the AFS sector, therefore we treat AFS as a competitive fringe. Studies into the profitability of AFS providers have found that their high fees per transaction tend to be offset with high marginal costs. These studies found that relatively low fixed costs of entry lead to a high level of competition in the AFS industry (Flannery and Samolyk 2005; Skiba and Tobacman 2007).

---

**10** For the sake of simplicity we are assuming that the bank can earn interest,  $r$ , instantaneously on the amount deposited by consumers. A more realistic setup would have consumers drawing down on their deposits continuously over the period, with the bank earning interest on an average deposit balance of  $\frac{1}{2}D_B$ . This adjustment would not substantively change any of our analysis. In effect  $r$  would become  $\frac{r}{2}$  throughout the rest of our analysis.

**11** One possible extension of our analysis would be to assume that consumers draw down a fixed amount of their deposits each period. In such a setup some low-income consumers would hit the "zero-bound" on their deposits, whereas higher income consumers would always be left with some positive balance in their accounts. This would make low-income consumers even less profitable for banks, and might lead them to charge penalty fees for consumers that reach the "zero-bound" (we observe these as overdraft fees in practice). For the sake of simplicity we bundle these potential fees in the general fee charged by the bank.



The bank takes the AFS fee,  $f_A$ , as given and equal to the constant marginal cost of providing AFS services,  $c_A$ . The bank chooses its customers by choosing  $f_B$ , which in effect determines  $w^*$ .<sup>12</sup> Differentiating (3) with respect to  $w^*$  we have:

$$\frac{\partial \pi_B}{\partial w^*} = -rw^*g(w^*) - (\theta w^* + c_A)g(w^*) + \theta [1 - G(w^*)] + c_B g(w^*) \quad (4)$$

The first and second terms on the right-hand side are the loss in interest revenue and fees from the marginal consumer at  $w^*$ . The third term is the gain from higher fees charged to all remaining bank customers. The final term is the cost savings from not providing services to the marginal consumer.

**Lemma 1.** *The second-order condition for profit maximization is satisfied for any income distribution with the non-decreasing hazard rate property.*<sup>13</sup>

*Proof.* See Appendix A. □

The proof for Lemma 1 shows that under the non-decreasing hazard rate assumption the second derivative of the bank's profit function is negative, and so the second-order condition for a maximum is satisfied.

From (4) and the assumption that the density of our income distribution is zero below  $a$ , we have that marginal profit is positive for any income below  $a$ . Therefore the bank will not charge a fee below the point where the consumer earning the lowest income will be indifferent between relying on the bank or AFS.<sup>14</sup> Using Eq. (2), this gives us a lower bound for the fee charged by the bank:

$$f_B \geq \theta a + c_A \quad (5)$$

Lowering the fee below this level will not add any new consumers and will cost the bank revenues from existing customers. Raising fees above this level would only be profitable if the right side of Eq. (4) is positive for the lowest income level,  $a$ . In addition, whether or not a bank prices itself out of the market depends on

---

<sup>12</sup> Note that if we had assumed a perfectly competitive banking market,  $f_B$  would be less than  $c_B$  in order to satisfy the zero profit condition. In addition to fees, banks earn revenues by investing consumer deposits. Therefore the zero profit condition for banks is a bit more complicated (see Freixas 2008, for a more detailed discussion). If banks only played the role of financial warehouses, then the investment return portion of the profit function would disappear and we would be in a more typical Shaked and Sutton (1982) setting.

<sup>13</sup> A non-decreasing hazard rate requires that the ratio  $\frac{g}{1-G}$  is not decreasing with  $w$ .

<sup>14</sup> In fact this is also true for cases where  $g(a) = 0$ . When the probability of earning the subsistence level of income is zero then marginal profit is positive at that income level. This result is significant for our condition for an interior solution.

the value of (4) at  $w^* = a + h$ . If the partial derivative of the profit function is negative at that income level, then the bank would have an incentive to lower its price to at least attract the wealthiest consumer in the market.

Evaluating the differential at these two points, i.e.  $w^* = a$  and  $w^* = a + h$ , gives us conditions on our model parameters that would allow for a bank to operate in a market but only target a portion of the consumer population:

$$(r + \theta)a - \frac{\theta}{g(a)} < c_B - c_A < (r + \theta)(a + h) \tag{6}$$

The right-hand side condition assures that it is worth it for a bank earning  $r$  and providing quality of service  $\theta$  to enter a market where the wealthiest consumers earn  $a + h$ . The left-hand side condition is when such a bank would not cater to the poorest consumers in the market, in other words it is the requirement for financial exclusion.

From the two conditions in (6) we can see that as long as there is enough income in a community the bank will choose to enter the market. In addition, if there is significant difference between the technology of the two types of financial service providers, as captured by  $c_B - c_A$ , relative to the income of the poorest consumer, then the left-hand condition in (6) holds and we have an interior solution where the bank targets a portion of the consumer population. These preliminary results seem to match what we would expect. Banks that are targeting higher-income consumers are more likely to provide better services in exchange for higher absolute fees, while financial companies targeting poorer neighborhoods are more likely to provide very basic services and charge lower fees. Figure 2 demonstrates the tradeoff from raising the cutoff level of income for the case of a uniform income distribution.

Interestingly, the condition for exclusion on the left-hand side of (6) is a weaker condition on the level of  $a$  than the requirement for profitability of the

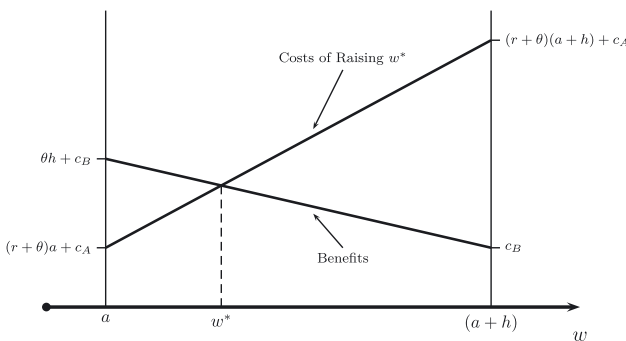


Figure 2: Costs and benefits of deposits.

poorest consumer,  $(r + \theta)a + c_A > c_B$ .<sup>15</sup> Therefore, the inability of the bank to price discriminate across consumers exasperates the extent of exclusion.<sup>16</sup>

In practice 2nd degree price discrimination is inherent in banks' pricing structure. Whether or not this would mean that our model underestimates or overestimates exclusion depends on the type of price discrimination that occurs in practice. Price discrimination in retail banking usually comes in two forms. The first is through indirect fees, the theoretical foregone interest consumers could earn if they invested their funds in a risk free asset rather than depositing them in a bank. Indirect fees are likely to be progressive for two reasons. Firstly, poorer consumers are less likely to have access to a risk-free outside option. It is also likely that higher income consumers have access to higher rates of returns on their investments and better outside options. Secondly, higher-income consumers on average have higher deposit balances, and so forego a greater amount of interest income. This form of discrimination is starker in the upper range of the income distribution, far away from  $w^*$ , and so is less relevant to our analysis in this paper.

The second form of price discrimination is in the direct fees charged by banks. These fees tend to be in the form of maintenance fees, overdraft fees and other fees that mainly apply to low-income and low-balance accounts. And so direct fees are a form of regressive price discrimination and are more pronounced at the lower end of the income distribution, closer to  $w^*$ . This second form of price discrimination is more relevant in our case as we are focused on the lower range of the income distribution. We are able to show that in the presence of this type of price discrimination our results regarding the impact of changes in the interest rate and the income distribution on exclusion would still apply.

## 2.3 Equilibrium

If the conditions for an interior solution are satisfied, then there is a profit maximizing level of  $w^*$  such that:

$$r + \theta = \frac{\theta[1 - G(w^*)]}{w^*g(w^*)} + \frac{c_B - c_A}{w^*}. \quad (7)$$

Whether or not the cutoff income for financial exclusion is decreasing with the rate of return available to the bank,  $r$ , depends on how the cumulative distribution function changes with  $w^*$ . If the first term on the right hand side of (7) is non-increasing with  $w^*$ , then we must have that an increase in the rate of

---

<sup>15</sup> This condition says that the poorest consumer is profitable if the bank sets fees such that  $w^* = a$ , that is such that the poorest consumer's participation constraint is binding.

<sup>16</sup> If the bank could perfectly price discriminate then it would set fees such that only the unprofitable consumers are excluded.

return available to the bank leads to lower cutoff level of income. As we show in Proposition 1 a non-decreasing hazard rate means that  $w^*$ , and in turn  $G(w^*)$ , are decreasing with  $r$ .<sup>17</sup>

**Proposition 1.** *If  $G(w)$  has a non-decreasing hazard rate then the cut-off level of income,  $w^*$ , and the percentage of unbanked,  $G(w^*)$ , are both decreasing with the interest rate,  $r$ .*

*Proof.* See Appendix A. □

The proof of Proposition 1 uses the non-decreasing hazard rate property of the income distribution to show that  $w^*$  is decreasing with  $r$ . Taking the differential of (4) with respect to  $r$  it is straightforward to show that a non-decreasing hazard rate is a sufficient but not necessary condition for  $\frac{dw^*}{dr}$  to be negative. The result for  $G(w^*)$  follows accordingly. We provide evidence for this result in our empirical analysis below.

The rate of return on deposits,  $r$ , is an important factor in the strategy of the bank. A higher  $r$  makes deposit assets more profitable for the bank, and less likely that poor consumers will be excluded. To the extent that exclusion from the financial sector negatively impacts low-income households, a higher return available for banks could be seen as a positive social outcome. This result is a bit misleading as we do not consider what factors determine  $r$ . Higher returns for banks can be due to greater risk and uncertainty in the bank's investment portfolio, which can be a negative outcome for the overall consumer population. This tradeoff became more apparent following the 2008 financial crisis and has spurred debate about the role of banks as deposit-taking institutions. It is not clear to what extent banks should focus purely on safeguarding consumer deposits rather than their rate of return on those deposits. Despite this tradeoff, Proposition 1 is a potential argument against the notion of limiting banks to only serving as money warehouses. If banks were not allowed to earn a return on customer deposits they would either respond by lowering the quality of deposit services,  $\theta$ , or more likely by raising fees, and in effect increasing financial exclusion.

---

<sup>17</sup> We are assuming that  $c_A$  and  $c_B$  are not impacted by changes in  $r$ . It is possible to argue that marginal costs to the AFS and the bank,  $c_A$  and  $c_B$ , may be affected by changes in the interest rate,  $r$ . Allowing for this would have an ambiguous impact on our results that depends on whether  $c_A$  or  $c_B$  is more sensitive to changes in  $r$ . If  $c_A$  is more sensitive then a rise in  $r$  would make AFS pricing less competitive relative to the banks. Banks would be able to raise fees in response to lower  $r$  without losing as many customers, so we would be overestimating our effects. The reverse would be true if  $c_B$  was more sensitive. We are not aware of any evidence that changes in the interest rate impact AFS and banks differently.

Another possible interpretation of  $r$  and its impact on exclusion is in the context of an economic recession. Zero or negative economic growth tends to coincide with periods of low returns on investments for financial institutions. Our model predicts that a low rate of return on deposits forces banks to increase direct fees on deposit customers, suggesting that financial exclusion is likely to increase in periods of slow to negative economic growth.<sup>18</sup>

## 2.4 Distribution of Income

In the context of our banking model  $a$  and  $h$  are measures of the standard of living and are defined relative to the technologies of the financial service providers,  $c_A$  and  $c_B$ . On their own they are not a sufficient summary statistics and only vaguely represent changes in dispersion. In this section we turn to a specific distribution function in order to more formally consider the impact of the distribution parameters of the consumer population on the percentage of unbanked. In particular, a unimodal distribution of income is considered a realistic representation of what we observe in practice, and so for our analysis we will work with a log-normal distribution of income.<sup>19</sup> As long as the log-normal satisfies our non-decreasing hazard rate assumption, then our results in (6) and (7) would still hold and there would be a unique profit maximizing value of  $w^*$ . Whether or not the log-normal has a non-decreasing hazard rate depends on the value of the parameters, mean log-income ( $\mu$ ) and standard deviation of log-income ( $\sigma$ ). In our empirical analysis we show that the log-normal is a good representation of the distribution of income in the regions we study. We also show that the second-order condition for a maximum is satisfied for all of the cases considered.

When income can be represented by a log-normal distribution the first-order condition in Eq. (4) can be rewritten as

$$\begin{aligned} \frac{\partial \pi_B}{\partial w^*} = & \left[ -(r + \theta) w^* + c_B - c_A \right] \left( \frac{e^{-\frac{(\ln w^* - \mu)^2}{2\sigma^2}}}{w^* \sigma \sqrt{2\pi}} \right) \\ & + \frac{\theta}{\sigma \sqrt{2\pi}} \int_{w^*}^{\infty} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx = 0 \end{aligned} \tag{8}$$

<sup>18</sup> In this instance, when considering the impact of a recession, we are ignoring any impact on the distribution of income. Clearly a recession might have redistributive effects or lead to a decrease in the standard of living, but here we are focusing only on the relation between periods of slow economic growth and the rate of return available to banks. We will consider the distributive impact on exclusion below.

<sup>19</sup> In Appendix B, we consider an alternative distribution function, the Pareto distribution.

We now consider how changes in the parameters of the log-normal distribution impact  $w^*$  and the proportion of unbanked,  $G(w^*)$ :

**Proposition 2.** *If income is log-normally distributed and if the condition for a non-decreasing hazard rate is satisfied:*

- *the proportion of unbanked,  $G(w^*)$ , is decreasing with  $\mu$ .*
- *$w^*$  is increasing with  $\mu$ ;*

*Proof.* See Appendix A. □

An increase in the mean of log-income without changes to the variance is approximately equivalent to a right-shift of the distribution. As we can see from Proposition 2 such a shift would result in an increase in the cut-off level of income, but would still lead to a lower percentage of unbanked as there would be fewer households earning an income below  $w^*$ . The intuition for this result comes from the bank's two sources of revenues, direct fees and indirect revenues earned on deposits. If all households in a neighbourhood earn a higher income the indirect revenue the bank earns per customer increases relative to direct fees. This would incentivise the bank to take on a higher percentage of households in order to increase its level of deposits.

**Proposition 3.** *If income is log-normally distributed, if the condition for a non-decreasing hazard rate is satisfied and if  $\ln w^* < \mu$ :*

- *the proportion of unbanked,  $G(w^*)$ , is increasing with  $\sigma$ ;*
- *if  $(\ln w^* - \mu)^2 > \sigma^2$ , then  $w^*$  is decreasing with  $\sigma$ .*

*Proof.* See Appendix A. □

This result demonstrates how greater income disparity, as measured by  $\sigma$ , may lead to a greater percentage of the population to be excluded from mainstream banking. An increase in  $\sigma$  without a change to  $\mu$  is a mean-preserving spread of income, resulting in a higher concentration of households in the tails of the distribution. The results in Proposition 3 suggest that in such a scenario, as long as  $w^*$  is not too high, it is likely that the bank will choose a lower  $w^*$  (by lowering fees) in order to pursue the households at the bottom of the distribution. But the decrease in  $w^*$  would not offset the increase in the concentration of households below  $w^*$ , resulting in a greater percentage of unbanked.

We now turn to the empirical section of our paper where we will use data on financial inclusion in the US in support of the hypotheses presented in Propositions 1, 2 and 3.

### 3 Data

In order to analyse the performance of the model built in the previous section vis-à-vis the actual data, we construct a balanced panel of 244 US geographic areas or markets, defined below, for the years 2009, 2011, 2013 and 2015, the only years for which information on access to banking services in the US is available. The final dataset includes data by geographic area and year on the proportion of unbanked,  $p$ , the corresponding cutoff level of income needed to access banking services,  $w^*$ , the interest rate margin,  $r$ , and the mean,  $\mu$ , and the standard deviation,  $\sigma$ , of the log of income. The dataset also includes information on single data points for the cost incurred by banks to open and manage an account,  $c_B$ , and the composite marginal cost to access AFS,  $c_A$ . Table 1 provides a summary statistics for some of these variables.

Information on whether households have a bank account and on their income are taken from the Current Population Survey (CPS), the standard household-level survey run in the US jointly by the US Bureau of Labor Statistics and the US Census Bureau. We make use of the Banked/Unbanked Supplement that the CPS conducted, in collaboration with the FDIC, in order to study access to banking services. This Supplement was run only in January 2009 and June 2011, 2013 and 2015, hence the years of our study. The main question of interest is “Do you or someone in the household have a bank account?” We construct the proportion of unbanked,  $p$ , as the weighted proportion of households by geographic area and year that respond negatively to this question, with weights provided by the CPS

**Table 1:** Summary statistics.

Variable	2009	2011	2013	2015	Total	
	Mean	Mean	Mean	Mean	Mean	St. dev.
Unbanked, $p$	0.080	0.085	0.080	0.074	0.080	0.050
Cutoff level of income, $w^*$	10,878	10,512	10,907	10,944	10,810	3062.0
Interest rate margin, $r$	0.046	0.043	0.037	0.036	0.041	0.008
Mean, log-income, $\mu$	10.596	10.558	10.611	10.647	10.603	0.251
Standard deviation, log-income, $\sigma$	0.915	0.934	0.928	0.921	0.924	0.104

The table reports the weighted mean for the proportion of unbanked,  $p$ , the cutoff level of income that makes consumers indifferent between holding a bank account or not,  $w^*$ , the interest rate margin,  $r$ , and the parameters of the log-normal distribution of income, i.e. mean,  $\mu$ , and standard deviation,  $\sigma$ , of log-income, based on 244 US geographic areas and separately for 2009, 2011, 2013 and 2015. The table also reports the weighted mean and standard deviation for each variable over the sample (columns under ‘Total’).

Supplement. Table 1 shows that the proportion of unbanked was approximately 8% in the first period analysed, increased between 2009 and 2011, and then decreased between 2011 and 2015. This variation will be the subject of a more in-depth analysis later.

The income data is taken from the same Banked/Unbanked Supplement. Individuals are asked how much the household they belong to earns annually based on a set of sixteen income bands. The actual level of income assigned to each household is the mid-point of each of these bands. Based on this household-level income data, we estimate the weighted mean,  $\mu$ , and standard deviation,  $\sigma$ , of the log of income, by geographic area and year, with weights provided by the CPS Supplement. Looking back at Table 1, we can see that average income first decreased between 2009 and 2011 and then jumped up over the next two periods. The opposite happened to the dispersion of income, with an initial increase followed by smaller decreases in the subsequent periods.

The household-level income data from the Banked/Unbanked Supplement is preferred to the data from the more standard CPS March Supplement because it allows us to match exactly the sample of households used to calculate the percentage of unbanked, even though the latter provides more detailed income data. Both the mean and the standard deviation of log-income by geographic area and year are significantly correlated between the two different datasets, with correlation coefficients of 0.71 and 0.25, respectively.

The theoretical model in the previous section makes use of the cutoff level of income,  $w^*$ , at which households are indifferent about holding a bank account. On an aggregate level (such as within a geographic area) and based on a given distribution of income, this cutoff level is directly linked to the proportion of unbanked. Therefore, we calculate this cutoff level using the inverse of the log-normal cumulative distribution function based on the values for the proportion of unbanked and the parameters of the log-normal distribution estimated for each geographic area and year.<sup>20</sup>

Regarding the geographic areas used to aggregate the household-level data, we also use the information provided by the CPS. The only geographic variables included in the CPS are states, Metropolitan Statistical Areas and, in some cases, Individual Central Cities linked to each Metropolitan Statistical Area. More disaggregated geographic information is not available due to the relatively small size of the sample included in the CPS. By combining these three geographic variables we can divide the US into 356 areas. In terms of our theoretical framework each

---

**20** It should be noted that in reality it can happen that households below that level of income hold a bank account and vice versa. For simplicity, this possibility is not taken into account in our model.



area represents the place where the bank has a local monopoly (or, equivalently, competes with a fixed number of other banks, as in Freixas 2008). Moreover, these geographic areas may vary considerably by size and population. However, this is not an issue because we will also provide estimations of the model using random and fixed effects, taking into account time-invariant characteristics of each geographic area. The final sample excludes geographic areas in which the proportion of unbanked is equal to zero in any given year. Thus, we arrive at a balanced sample of 976 observations, i.e. 244 geographic areas times four years.<sup>21</sup>

The remaining variables are calculated from various sources, but mainly using data provided by the FDIC. In particular, the interest rate margin,  $r$ , is calculated at the state level as the total interest income earned by commercial banks divided by the average total earning assets over a given year, all information provided by the FDIC and not readily available at a more disaggregated level.<sup>22</sup> We calculate the interest rate margin for all years between 2009 and 2015. In our counterfactual analysis we use these values to study the impact of the recent financial crisis on access to banking services. During the 2009–2015 period, the interest rate margin decreased sharply from 4.6% down to 3.6%.

Following the theoretical framework, the marginal cost of AFS,  $c_A$ , is equal to the fees paid by AFS customers. To estimate the marginal cost of an AFS provider we build a composite annual cost of AFS use by an average unbanked household. Average use is calculated by estimating the number of times a typical unbanked household uses each one of the four main AFS transaction services (check cashing, pre-paid cards, money orders and remittances) and multiplying this by an estimated average fee associated with each use of each product. The composite annual cost is then given by the sum over the estimated annual fees of the four products. We estimate the typical use based on the survey results of the CPS Banked/Unbanked June 2013 Supplement. The survey asks households whether they have used each of the four products according to three potential frequencies: ever, in the last 12 months but not including this past month, or this past month. While for the first category, it is obvious that households have never used a particular service, we need to make assumptions about the other

---

**21** As a robustness check, we estimate the model with an unbalanced panel that excludes only observations with a proportion of unbanked equal to zero in that given year. This gives us a panel with 1221 observations, divided in 343 geographic areas. Moreover, we estimate the model changing the likelihood function given below to account for the fact that we only include observations with positive proportions of unbanked in all years, which implies using a conditional likelihood function. All results carry through and are available upon request.

**22** In the calculation of the interest rate margin we subtract the national percentage interest paid on domestic non-jumbo checking accounts, but it should be noticed that this has been very close to zero throughout the period considered and its inclusion does not affect the results.

two categories. If households have used a product in the last year, but not in the last month, we assume that they use it twice a year. If they have used a product in the last month we assume that they use it six times a year. With regards to the cost of every use of each product, we use information from the websites of major AFS providers. Using this approach our estimates for the marginal cost of AFS range from 22 to 59 dollars, with the average value of 40 dollars being the preferred choice. This value is used for all observations.

The marginal cost of opening and managing an account by a bank,  $c_B$ , is estimated to be equal to 51 dollars, with a range from 44 to 54 dollars, based on Haberfeld (2002) and confidential information provided by a major US banking institution. This value is assumed to be the same for all observations.

We also include data on competition, measured by the number of branches at the state level, and market concentration, measured by the share of deposits for the top five banks at the state level, to be used in our reduced-form estimation. All information is taken from the FDIC and more disaggregated data are not readily available.

Finally, we make use of data on the percentages of prisoners and homeless people in the reduced-form analysis to take into account the fact that the CPS excludes these people from its sample and they are likely to be at severe risk of financial exclusion. These data are calculated at the state level and are taken from the Bureau of Justice Statistics and the US Department of Housing and Urban Development respectively. The percentages of prisoners and homeless people have been declining during the period 2009–2015.

## 4 Reduced-form Estimation

The correlation matrix in Table 2 provides some informal evidence as to the relationship between access to banking services, the interest rate margin and the distribution of income. The table confirms the basic suggestions of the model. There is a negative and statistically significant relationship between the proportion of unbanked and the mean level of income. Also, there is a positive and statistically significant relationship between the proportion of unbanked and the standard deviation of the log-normal distribution of income, meaning that in areas where incomes are more spread out (and, thus, there is potentially more inequality) the proportion of unbanked is higher. However, the interest rate margin does not seem to be statistically correlated with the other variables, but it is important to notice that this is an unconditional correlation.

We can further examine the basic relationships outlined in the model above by linearly regressing the proportion of unbanked on a set of relevant variables.

**Table 2:** Correlation matrix.

	Unbanked, $p$	Cutoff, $w^*$	Interest, $r$	Mean, $\mu$	St. dev., $\sigma$
Unbanked, $p$	1.00				
Cutoff level of income, $w^*$	0.55***	1.00			
Interest rate margin, $r$	0.03	0.06	1.00		
Mean, log-income, $\mu$	-0.53***	0.08***	-0.04	1.00	
Standard deviation, log-income, $\sigma$	0.32***	-0.12***	-0.04	-0.46***	1.00

The number of observations is 976. \*\*\*indicates coefficients significantly different from zero at 1% level.

Among the regressors, we include the interest rate margin, the mean and standard deviation of log-income, the number of branches, the share of deposits of the top five banks, the proportion of unemployed, average years of education, the proportion of the population who have passed the General Education Development (GED) tests, the proportion of different ethnicities and races (non-Hispanic White, Hispanic, Black and Asian), and the percentages of prisoners and homeless people. We use the following equation:

$$p_{it} = \beta_0 + \beta_1 r_{it} + \beta_2 \mu_{it} + \beta_3 \sigma_{it} + \beta_4 X_{it} + \phi_i + \psi_t + \varepsilon_{it} \tag{9}$$

where  $X$  is the set of control variables excluding the interest rate margin, the mean and the standard deviation of income,  $\phi$  represents the geographic areas fixed effects,  $\psi$  represents the time fixed effects,  $\varepsilon$  is a classical error term, and the subscripts  $i$  and  $t$  represent geographic areas and years respectively. Table 3 shows the results of these regressions. The three columns differ in the sample used. The regression in column (1) includes all observations, including those with the proportion of unbanked equal to zero. The regression in column (2) excludes those observations with the proportion of unbanked equal to zero in that given year, thus creating an unbalanced panel. Last, the regression in column (3) excludes all observations with the proportion of unbanked equal to zero in any year, thus creating a balanced panel. All regressions include fixed effects for geographic areas and time periods and are significant overall, with R-squared values ranging from 0.25 to 0.27.

The above regressions tell us that there is a statistically negative relationship between the interest rate margin and the proportion of unbanked. We also find a negative relationship between the mean of log-income and the proportion of unbanked and a positive relationship between the standard deviation of log-income and the proportion of unbanked. All these results are robust and carry through in all of these regressions, independently of the sample used. Contrary to the results in Bord (2017) and Celerier and Matray (2019), there is no relationship

**Table 3:** The reduced-form determinants of unbanked, Eq. (9).

	(1)	(2)	(3)
Interest rate margin, $r$	-0.44* (0.24)	-0.58*** (0.21)	-0.68*** (0.22)
Mean, log-income, $\mu$	-0.06*** (0.01)	-0.05*** (0.02)	-0.06*** (0.02)
Standard deviation, log-income, $\sigma$	0.05*** (0.01)	0.06*** (0.02)	0.05*** (0.02)
No branches	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)
Bank concentration (top 5)	0.06 (0.04)	0.06 (0.04)	0.06 (0.04)
Unemployed	0.11** (0.05)	0.10* (0.05)	0.08 (0.06)
Years education	0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)
GED	-0.12*** (0.03)	-0.09** (0.04)	-0.10** (0.05)
White	-0.13** (0.06)	-0.14** (0.07)	-0.13** (0.07)
Hispanic	-0.04 (0.07)	-0.05 (0.08)	-0.06 (0.09)
Black	0.04 (0.06)	0.03 (0.07)	0.03 (0.08)
Asian	-0.12* (0.07)	-0.09 (0.08)	-0.07 (0.10)
Prisoners	0.08* (0.04)	0.07 (0.04)	0.06 (0.05)
Homeless	-0.02 (0.04)	0.00 (0.04)	0.01 (0.04)
Geographic areas f.e.	Yes	Yes	Yes
Year f.e.	Yes	Yes	Yes
No of obs.	1424	1221	976
$R^2$ (within)	0.26	0.25	0.27
$F$ statistic	16.49	12.50	11.95

The dependent variable is the proportion of unbanked by geographic area and year. All regressions are weighted with weights provided by the CPS. Standard errors are shown in parentheses. \*, \*\* and \*\*\* indicate coefficients significantly different from zero at 10, 5 and 1% level respectively.

between competition (number of bank branches) and market concentration (share of deposits of top five banks) on the one hand and the proportion of unbanked on the other. These results confirm the important role that the interest rate margin and the distribution of income play as determinants of financial exclusion, as

suggested by our model. With regards to the other variables, we find that areas with a greater proportion of the population with at least a GED tend to have fewer unbanked households and white Americans are estimated to be more likely to have a bank account. This suggests that lack of information (also due to lack of education) and racial discrimination may play a role in financial exclusion. On the other hand, controlling for all other factors, unemployment does not affect financial exclusion. Both interesting results to keep in mind for future work. Finally, the coefficients on the percentages of prisoners and homeless people are not statistically significant and, thus, their exclusion from CPS does not seem to cause a bias in our main results.

## 5 Structural Estimation

We now structurally estimate our model, rather than via reduced form. Moving on to a structural estimation is important in this case. It allows us to see how well the model matches the actual data both on aggregate and on a market-by-market basis. It also makes it possible to run a set of counterfactual scenarios to study the quantitative effects of changes in the variables of interest on the proportion of unbanked. This can be used, for instance, to analyse whether changes in mean income are more or less important than changes in the dispersion of income for financial exclusion.

### 5.1 Econometric Model

We start from the stochastic specification used in order to estimate the first-order condition in Eq. (8). We assume an additively separable error term, such that the estimating equation is given by

$$\left[ -(r + \theta) w^* + c_B - c_A \right] \left( \frac{e^{-\frac{(\ln w^* - \mu)^2}{2\sigma^2}}}{w^* \sigma \sqrt{2\pi}} \right) + \frac{\theta}{\sigma \sqrt{2\pi}} \int_{w^*}^{\infty} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx - u = 0 \quad (10)$$

where  $u$  is the error term. Besides the error term, the only parameter to be estimated in Eq. (10) is  $\theta$ , the benefit of having a retail bank account. This is identified via variation in the interest rate margin, mean income and income dispersion across markets and years. All these variables are assumed to be exogenous with respect to the proportion of unbanked at this level of aggregation, at least once geographic area fixed effects are taken into account. While it is expected that parameter  $\theta$  takes on a positive value, the benefit of having a retail bank account for consumers earning  $w^*$  can potentially turn out to be negative. This depends, among other

things, on the difference between the marginal cost of a formal bank account and the marginal cost of AFS.

We specify five different versions of the stochastic term  $u$ , which correspond to the standard ones used in empirical panel data analysis: pooled data (POOL), one-way random effects (RE1), one-way fixed effects (FE1), two-way random effects (RE2) and two-way fixed effects (FE2). All five versions of the stochastic model in Eq. (10) are estimated via maximum likelihood.<sup>23</sup>

## 5.2 Estimation Results

Table C1 in Appendix C presents all parameter estimates.<sup>24</sup> We focus on the two-way fixed effects model given that the likelihood ratio (LR) tests show that this is the preferred specification.

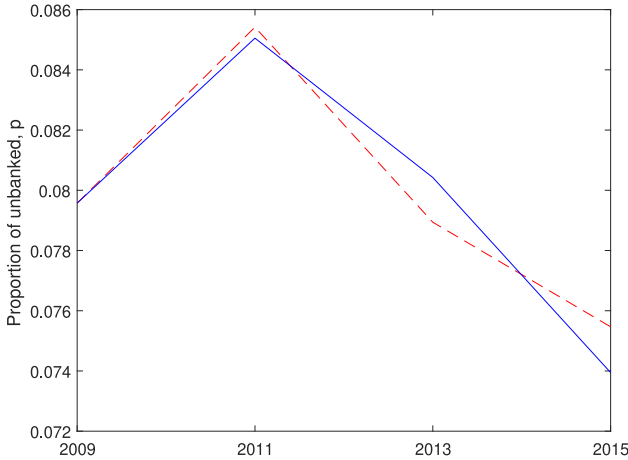
The results show that the estimated benefit of having a retail bank account,  $\theta$ , does not vary significantly across the five specifications and it is equal to 0.009. In all five cases,  $\theta$  is significantly different from zero at the 1% level based on the bias corrected and accelerated clustered bootstrap confidence intervals. The size of the coefficient implies that banks provide a benefit of slightly less than 1% to consumers who use their services relative to AFS. Also, based on the size of  $\theta$  and the expression for bank charges,  $f_B$ , given in Eq. (2), we can calculate that banks charge 137 dollars per year for their services to the average customer across the US. This is equivalent to about 11 dollars per month and is close to figures estimated by industry analysts (Bakker et al. 2014).

We now study how well the two-way fixed effects model performs relative to the actual data. In particular, we are interested in how well the model does in terms of explaining changes in the percentage of unbanked. Thus, we let all relevant variables (i.e. interest rate margin, mean and standard deviation of income) change in accordance with the actual data. We also vary the fixed effects based on the estimated values, while keeping the idiosyncratic error terms at the values estimated for 2009. We can then compare the predicted changes in the percentage of unbanked from our model with the changes in the actual data.

Figure 3 compares the actual data and the model fit for the proportion of unbanked when the interest rate margin and the parameters of the income

<sup>23</sup> In order to avoid the well-known incidental parameters issue that occurs when a fixed effects model is estimated via maximum likelihood, we eliminate the fixed effects by writing the model in differences from means for each geographic area and year.

<sup>24</sup> It is important to note that in all our estimations the second derivative of the profit function with respect to  $w^*$  is calculated to be negative for all observations (see the proof for Lemma 1 in Appendix A for the functional form of the second-order condition). Thus, all estimations meet the assumption of the model with respect to the concavity of the profit function.



**Figure 3:** Model versus data: changes in the interest rate and the distribution of income. The blue solid line represents the actual data and the red dashed line the model fit.

distribution change and the time-variant unobservable factors are also taken into account. The model can replicate closely the increase in the proportion of unbanked in the period 2009–2011 as well as its subsequent decrease over the period 2011–2015. While for the period 2011–2013 the model slightly overpredicts the real extent of the decrease in the proportion of unbanked, in the period 2013–2015 the model slightly underpredicts its decrease. However, the differences between the actual data and model predictions on average are not statistically significant.<sup>25</sup>

Table 4 provides a deeper look at how well the model fits the actual data. For each time period the direction of change in the proportion of unbanked observed in the data is compared with the direction predicted by the model (based on 99% confidence intervals for the estimated parameters). If the model could predict perfectly the direction of the change in the proportion of unbanked, all observations would fall either in the top-left or bottom-right cells for each period and the sum of the values in these two cells would be equal to one. We can see that the direction of change in the proportion of unbanked is predicted correctly by our model for at least 71% of observations (2011–2013 period) and for

<sup>25</sup> The small gap between the actual data and the model fit in the 2011–2015 period is potentially because that the 2015 CPS data uses a revised set of Metropolitan Statistical Area delineations based on data from the 2010 Census. Consequently, some geographic areas may differ between 2015 and the earlier periods considered because of changes in geographic boundaries (FDIC 2015). Indeed, when we re-estimate our model including only those geographic areas for which the geographic boundaries do not change following the revised delineations, the slight gap between the actual data and the model fit that appears between 2011 and 2015 almost disappears. All the main results carry through and are available upon request.

**Table 4:** Model fit versus real data: direction of change in the proportion of unbanked.

		Real Data					
		2009–2011		2011–2013		2013–2015	
		Decrease	Increase	Decrease	Increase	Decrease	Increase
Prediction	Decrease	0.26	0.06	0.45	0.17	0.39	0.12
	Increase	0.16	0.52	0.12	0.26	0.16	0.33

The table shows the proportions of observations in each period based on the direction of change in the proportion of unbanked (i.e. decrease or increase) comparing the real data (columns) with the model predictions (rows).

up to 78% of observations (2009–2011 period). Therefore, we can claim that the model does a good job at predicting the direction of change in the proportion of unbanked for most observations.

## 6 Counterfactual Analysis

In this section we quantify the effects of changes in the interest rate margin (see Proposition 1), mean log-income (see Proposition 2) and the standard deviation of log-income (see Proposition 3) on the proportion of unbanked. We use the estimates from the two-way fixed effects model to conduct three counterfactual exercises, one for each proposition and related variable of interest. The counterfactuals are run by letting all variables and error terms change over time with the exception of the variable under analysis. Thus, in the figures below, the effects due to changes in the variable of interest are given by the difference between the line representing the actual data and that representing the model prediction.<sup>26</sup>

### 6.1 Changes in the Interest Rate Margin

Under the first counterfactual we keep the interest rate margin at its 2009 level for each geographic area and allow all other variables and error terms to change based on the actual data and our estimates. This allows us to see how much of the changes in the proportion of unbanked between 2009 and 2015 were due to changes in the interest rate margin, and so due to the financial crisis. This is based on the observation that one of the consequences of the financial crisis was

<sup>26</sup> If the line representing the actual data, i.e. the solid blue one, is above the line representing the model fit, i.e. the dashed red one, then the effect is positive and vice versa.



a considerable decrease in the interest rate margin, which according to our model should have had an effect on access to banking services.

Figure 4 shows how the proportion of unbanked changes over the 2009–2015 period in the actual data (solid blue line) and according to the first counterfactual (dashed red line) as a weighted average across all geographic areas in the sample. The model predicts that, without changes in the interest rate margin over this period, the proportion of unbanked would have been significantly lower, meaning that fewer households would have been priced-out of formal banking services. The reason is that when the interest rate margin decreases, as it happens during a financial crisis, banks can make fewer profits from any given level of deposits. Customers with small deposits become even more unprofitable and, thus, are pushed out of formal banking services. It is therefore possible to argue that the last financial crisis had a large and negative impact on financial inclusion via a decrease in banks' interest rate margins.

Table 5 makes explicit the quantitative effects of changes in the interest rate margin on the proportion of unbanked during this period. It shows how much an increase (decrease) in the interest rate margin is associated with a decrease (increase) in the proportion of unbanked by splitting the sample of geographic areas. While the large majority of geographic areas in our sample experienced substantial decreases in percentage terms in the interest rate margin, between  $-6.5\%$  and  $-12\%$ , few geographic areas went through equally large increases in the interest rate margin. These changes are associated with large percentage changes, of opposite sign but approximately in the same order of magnitude, in the proportion of unbanked.



**Figure 4:** Counterfactual 1, changes in the interest rate margin. The blue solid line represents the actual data and the red dashed line the model fit.

**Table 5:** Quantitative effects of changes in the interest rate margin.

	2009–2011		2011–2013		2013–2015	
	% $\Delta r$	% $\Delta p$	% $\Delta r$	% $\Delta p$	% $\Delta r$	% $\Delta p$
$\Delta r > 0$	47.687	–29.710	10.228	–8.466	3.086	–3.294
$\Delta r < 0$	–8.976	10.064	–12.317	14.381	–6.533	7.211

The first row represents changes in  $r$  and  $p$  for all geographic areas that observed an increase in the interest rate margin, separately for the periods 2009–2011, 2011–2013 and 2013–2015, while the second row represents changes in  $r$  and  $p$  for all geographic areas that observed a decrease in the interest rate margin. All values are in percentage changes.

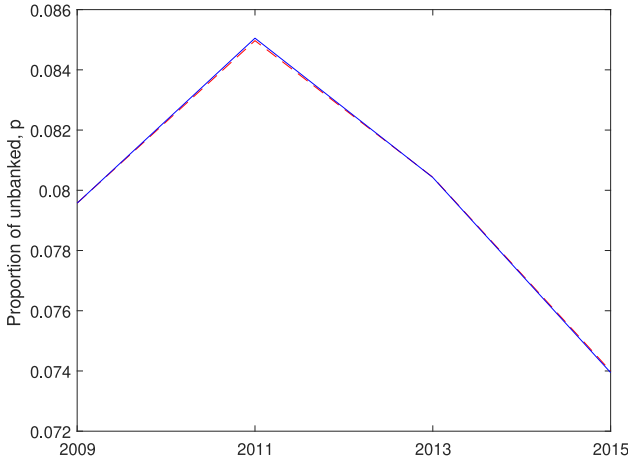
This result is relevant to the growing literature on the potential impact of monetary policy on the distribution of income and wealth (Amaral 2017; Lenza and Slacalek 2021; O’Farrell and Rawdanowicz 2017). The standard approach has been to look at the impact of monetary policy on asset prices, capturing the direct impact on wealth inequality. Our paper suggests that the second order effects may be just as important, monetary policy impacts interest rates, which in turn impacts access to banking by low-income households. This result is supported by Ampudia et al. (2018). They show that indirect channels are very important when looking at the impact of monetary policy on inequality. This is especially true when looking at households with a low level of assets.

## 6.2 Changes in Mean Income

The second scenario looks at the effect of changes in the average level of incomes on access to banking. For each geographic area we keep the mean log-income at the 2009 level and then we see how the proportion of unbanked changes accordingly as we let all other variables and error terms change. During this period the average level of income first decreases slightly, following the financial crisis, but then jumps up to a higher level than the initial one.

Figure 5 shows the weighted average values for the proportion of unbanked based on the actual data and the proposed counterfactual. According to our model, if mean income had not changed over this period of time there would be little difference in the proportion of unbanked.

These results suggest that changes in mean income have little effect on the proportion of unbanked, at least on an aggregate level. However, in order to properly quantify the mechanism under consideration, we disaggregate the average effects by splitting the sample of geographic areas into those that observe an increase in mean income and those that observe a decrease in mean income in



**Figure 5:** Counterfactual 2, changes in mean income. The blue solid line represents the actual data and the red dashed line the model fit.

each period. Table 6 shows that not only the percentage changes in mean income are relatively small, about 1% in all cases, but the effects on the proportion of unbanked are also rather small, usually less than one third of the size of the changes in mean income.

### 6.3 Changes in the Standard Deviation of Income

The third scenario looks at the effect of changes in the dispersion of income on access to banking. For each geographic area we keep the standard deviation of log-income at the 2009 values and then we see how the proportion of unbanked changes as all other variables and error terms change. During this period, income dispersion increases in the first period due to the crisis and then decreases by a smaller amount on average in the second and third periods.

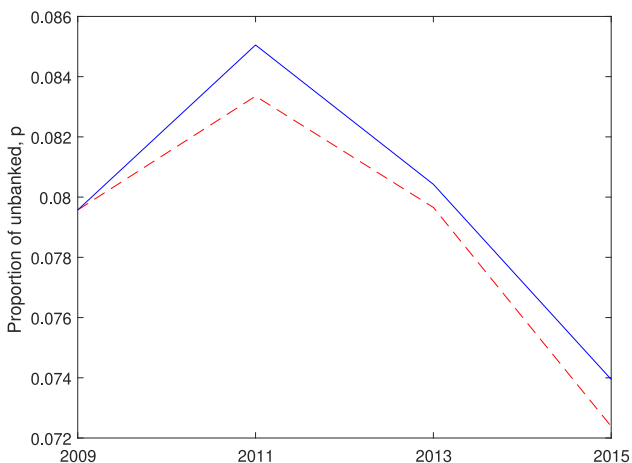
**Table 6:** Quantitative effects of changes in mean income.

	2009–2011		2011–2013		2013–2015	
	% $\Delta\mu$	% $\Delta p$	% $\Delta\mu$	% $\Delta p$	% $\Delta\mu$	% $\Delta p$
$\Delta\mu > 0$	0.910	-0.300	1.000	-0.251	1.207	-0.298
$\Delta\mu < 0$	-0.971	0.256	-0.731	0.217	-1.040	0.275

The first row represents changes in  $\mu$  and  $p$  for all geographic areas that observed an increase in mean income, separately for the periods 2009–2011, 2011–2013 and 2013–2015, while the second row represents changes in  $\mu$  and  $p$  for all geographic areas that observed a decrease in mean income. All values are in percentage changes.

Figure 6 shows the weighted average values for the proportion of unbanked based on the actual data and the model fit under this third counterfactual scenario. As per the predictions of our model, without changes in the standard deviation of income the proportion of unbanked would have been lower throughout the period, especially in 2011.

Finally, as done previously, Table 7 shows these effects in percentage terms depending on whether income dispersion increases or decreases in each geographic area and period. The effect in percentage terms of changes in the dispersion of income on the proportion of unbanked is much larger than for the case of changes in mean income. Thus, it is possible to argue that the negative distributional effects following from the last financial crisis on financial inclusion are more likely to be due to the increase in the dispersion of income rather than due to the decrease in mean income. More generally, exclusion from formal



**Figure 6:** Counterfactual 3, changes in the standard deviation of income. The blue solid line represents the actual data and the red dashed line the model fit.

**Table 7:** Quantitative effects of changes in the standard deviation of income.

	2009–2011		2011–2013		2013–2015	
	$\% \Delta \sigma$	$\% \Delta p$	$\% \Delta \sigma$	$\% \Delta p$	$\% \Delta \sigma$	$\% \Delta p$
$\Delta \sigma > 0$	8.981	13.484	7.681	11.573	9.850	13.984
$\Delta \sigma < 0$	-8.061	-11.818	-8.552	-9.884	-9.895	-11.325

The first row represents changes in  $\sigma$  and  $p$  for all geographic areas that observed an increase in the standard deviation of income, separately for the periods 2009–2011, 2011–2013 and 2013–2015, while the second row represents changes in  $\sigma$  and  $p$  for all geographic areas that observed a decrease in the standard deviation of income. All values are in percentage changes.

banking services is more affected by changes in the dispersion of incomes than by changes in their average levels.

## 7 Conclusion

In this paper we have looked to more formally analyze the determinants of exclusion from mainstream banking. We have used a supply-side model of banking services to demonstrate how under certain circumstances it might be optimal for the bank to exclude the lower income portion of the population. The model focuses on the interest rate margin and the distribution of income as fundamental determinants of access to retail banking.

Using localized US household data from 2009, 2011, 2013 and 2015 and a reduced-form regression analysis, we have tested our model's predictions. We have found evidence in favour of the fact that decreases in the interest rate margin and increases in income disparities lead to increases in the proportion of unbanked and higher financial exclusion.

Then, we have structurally estimated the model in order to quantify the effects of the interest rate margin and the distribution of income on access to banking. We have also used our structural estimation in order to show that banks provide a small but significant benefit to low-income consumers relative to relying on AFS providers, and so we are able to demonstrate a cost to financial exclusion.

The rate of return available to the bank,  $r$ , is found to play a positive role in reducing exclusion from mainstream banking. Allowing banks to invest customer deposits has a positive impact on the consumer population by reducing the direct fees they have to pay for banking services. To the extent that consumers do not have access to a risk free rate of return for their assets, these direct fees make up a big chunk of the costs of banking. By allowing the bank to reduce direct fees, a higher rate of return on deposits reduces exclusion from banking services, as well as increasing consumer surplus. But the positive impact of  $r$  depends on what drives the increase in returns for the bank. If an increase in  $r$  is associated with economic growth and better investment opportunities, then it can be seen as a win-win outcome for consumers and the overall economy. On the other hand, if increases in  $r$  are driven by higher risk in the bank's investment portfolio, the positive impact on consumers can be short lived; a phenomenon that we observed directly in the 2008 financial crisis. Future work on this topic should consider the tradeoff a bank faces when it chooses  $r$ , and how its choice of risk in its investment portfolio depends on the consumer population and the economic environment.

These results are a good demonstration of how introducing a heterogeneous consumer population adds greater depth to economic analysis. As far as we know, models of banking services have mainly ignored the role of income distribution in considering the strategic decisions of financial institutions.

In addition, there are some important policy implications to our model and empirical results. Our paper demonstrates that monetary policy can have second-order impacts on low-income consumers. Low interest rate policy in the face of a recession may lead to greater financial exclusion. As we have argued above, this would have a disproportionate impact on lower-income consumers, who are likely to be those that are most vulnerable to the negative consequences of the recession.

Similarly, greater banking regulation that reduces the rate of return earned by banks (or the interest rate charged by banks) could lead to banks raising fees, and so excluding a larger portion of the consumer population.

It is difficult to avoid loose monetary policy and banking regulation, our experiences over the last two decades has made that clear. But our results suggest that policy makers should seriously take into account these two second-order effects when setting policy and when considering costs and benefits to changes in regulation. The most effective approach would be long-term policy to mitigate the number of unbanked in the population generally, removing the concern of these second-order effects.

**Acknowledgments:** We thank participants at the RES Conference and at seminars at the University of Trento. All errors are ours. Declarations of interest: none.

## Appendix A: Proofs

*Proof of Lemma 1.* Differentiating (4) with respect to  $w^*$  we have the following second-order condition for profit maximization:

$$g(w^*)(r + 2\theta) > g'(w^*) [c_B - c_A - (r + \theta)w^*] \quad (\text{A1})$$

Using the identity in (4) we need:

$$\frac{r}{\theta} + 2 + \frac{g'(w^*)}{g(w^*)^2} [1 - G(w^*)] > 0 \quad (\text{A2})$$

Where we know that the final inequality holds as our non-decreasing hazard rate condition requires that  $\frac{g'(w^*)}{g(w^*)^2} [1 - G(w^*)] \geq -1$  and we have that  $\frac{r}{\theta} + 2 > 1$  by definition, giving us the required result.  $\square$

*Proof of Proposition 1.* The bank's choice of  $w^*$  is determined by the F.O.C. in (4). Differentiating (4) with respect to  $r$  and solving for  $\frac{dw^*}{dr}$  we have:

$$\frac{dw^*}{dr} = \frac{-w^*g(w^*)}{g(w^*)(r + 2\theta) + g'(w^*)[(r + \theta)w^* - (c_B - c_A)]} \quad (A3)$$

The numerator of (A3) is negative. To show that  $w^*$  is decreasing with  $r$  we need to show that the denominator is positive. Substituting in the identity from (4) we need:

$$\begin{aligned} g(w^*)(r + 2\theta) + \frac{g'(w^*)}{g(w^*)}\theta [1 - G(w^*)] &> 0 \\ \Rightarrow \frac{r}{\theta} + 2 + \frac{g'(w^*)}{g(w^*)^2} [1 - G(w^*)] &> 0 \end{aligned}$$

Which we have shown to hold in the proof for Lemma 1. Therefore we have that  $\frac{dw^*}{dr} < 0$  as required.

To show that the percentage of unbanked is decreasing with  $r$  we need:

$$\frac{dG(w^*)}{dr} = g(w^*)\frac{dw^*}{dr} < 0$$

Where we know that the inequality holds as our result above shows that  $\frac{dw^*}{dr} < 0$ .  $\square$

*Proof of Proposition 2.* To show the first part of the proposition we begin by noting that when using a log-normal distribution:

$$\frac{\partial g(w^*)}{\partial w^*} = g'(w^*) = -\frac{g(w^*)}{w^*} \left( 1 + \frac{\ln w^* - \mu}{\sigma^2} \right) \quad (A4)$$

Differentiating (8) with respect to  $\mu$  and using the identities from the F.O.C. in (4) and from (A4) we have:

$$\frac{dw^*}{d\mu} = \frac{\frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)}{\sigma^2} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx - \frac{(\ln w^* - \mu)}{\sigma^2} (1 - G(w^*))}{g(w^*) \left( \frac{r}{\theta} + 2 \right) + (1 - G(w^*)) \frac{g'(w^*)}{g(w^*)}} \quad (A5)$$

From the proof for Proposition 1 we have that our non-decreasing hazard rate assumption assures us that the denominator of (A5) is positive. To show that  $w^*$

is increasing with  $\mu$  we need to show that the numerator is positive. Noting that  $(\ln x - \mu)$  is a monotonically increasing function of  $x$  we have that:

$$\begin{aligned} \frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)}{\sigma^2} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx &> \frac{(\ln w^* - \mu)}{\sigma^2} \frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx \\ &= \frac{(\ln w^* - \mu)}{\sigma^2} (1 - G(w^*)) \end{aligned}$$

Therefore the numerator is also positive and we have that  $\frac{dw^*}{d\mu} > 0$  as required.

For the second part of the proposition we differentiate  $G(w^*)$  with respect to  $\mu$  to show:

$$\begin{aligned} \frac{dG(w^*)}{d\mu} &= \frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)}{\sigma^2} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx + \frac{dw^*}{d\mu} g(w^*) < 0 \\ &\Rightarrow -\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\ln w^* - \mu)^2}{2\sigma^2}} + \frac{dw^*}{d\mu} g(w^*) < 0 \\ &\Rightarrow \frac{dw^*}{d\mu} < w^* \\ &\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)}{\sigma^2} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx - \frac{(\ln w^* - \mu)}{\sigma^2} (1 - G(w^*)) \\ &< w^* g(w^*) \left( \frac{r}{\theta} + 2 \right) - \left( 1 + \frac{\ln w^* - \mu}{\sigma^2} \right) (1 - G(w^*)) \\ &\Rightarrow \theta w^* g(w^*) < w^* g(w^*) (r + 2\theta) - g(w^*) [(r + \theta)w^* - (c_B - c_A)] \\ &\Rightarrow 0 < g(w^*) (c_B - c_A) \end{aligned}$$

Where the last step is true by definition, giving us that  $\frac{dG(w^*)}{d\mu} < 0$  as needed.  $\square$

*Proof of Proposition 3.* We begin by proving the second part of the proposition. Differentiating (8) with respect to  $\sigma$  and using the identities from the F.O.C. in (4) and from (A4) we have:

$$\frac{dw^*}{d\sigma} = \frac{\frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)^2}{\sigma^3} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx - \frac{(\ln w^* - \mu)^2}{\sigma^3} (1 - G(w^*))}{g(w^*) \left( \frac{r}{\theta} + 2 \right) + (1 - G(w^*)) \frac{g'(w^*)}{g(w^*)}} \tag{A6}$$

From the proof for Proposition 1 we have that our non-decreasing hazard rate assumption assures us that the denominator of (A6) is positive. To show that  $w^*$



is decreasing with  $\sigma$  we need to show that the numerator is negative.

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{w^*}^{\infty} \frac{(\ln x - \mu)^2}{\sigma^3} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx < \frac{(\ln w^* - \mu)^2}{\sigma^3} (1 - G(w^*))$$

Using integration by parts we need:

$$(\ln w^* - \mu)w^*g(w^*) < (1 - G(w^*)) \left( \frac{(\ln w^* - \mu)^2}{\sigma^2} - 1 \right)$$

We have that the left side of the inequality is negative ( $\ln w^* < \mu$ ) and the right side is positive ( $(\ln w^* - \mu)^2 > \sigma^2$ ) by assumption, giving us that  $\frac{dw^*}{d\sigma} < 0$  as required.

For the first part of the proposition we need to show that  $\frac{dG(w^*)}{d\sigma} > 0$ .

$$\frac{dG(w^*)}{d\sigma} = -\frac{1}{\sigma}G(w^*) + \frac{1}{\sigma\sqrt{2\pi}} \int_0^{w^*} \frac{(\ln x - \mu)^2}{\sigma^3} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx + \frac{dw^*}{d\sigma}g(w^*) > 0$$

Using integration by parts this condition reduces to:

$$\begin{aligned} \frac{dw^*}{d\sigma} &> \frac{\ln w^* - \mu}{\sigma} w^* \\ \Rightarrow \theta(1 - G(w^*))[1 + (\ln w^* - \mu)] &> (\ln w^* - \mu)g(w^*)(r + \theta)w^* \\ &\Rightarrow \theta(1 - G(w^*)) > (\ln w^* - \mu)g(w^*)(c_B - c_A) \end{aligned}$$

Where the last inequality holds as we assume that  $\ln w^* < \mu$ . Note that we used integration by parts and the identity from the F.O.C. in (4) for the last two lines.  $\square$

## Appendix B: Pareto distribution

In this section, we consider an alternative distribution function, the Pareto distribution, where income is greater than or equal to our lower bound  $a$  (this is equivalent to  $h = \infty$ ). In this case, the cumulative distribution and density functions are given by:

$$G(w) = 1 - \left(\frac{a}{w}\right)^\alpha \quad g(w) = \frac{\alpha}{a} \left(\frac{a}{w}\right)^{\alpha+1} \quad \text{s.t. } \alpha > 1$$

Under this distribution  $g(a) = \frac{\alpha}{a}$ , where  $\alpha$  is a shape parameter of the distribution. Therefore condition (6) becomes:<sup>27</sup>

$$(r + \theta)a - \frac{\theta a}{\alpha} < c_B - c_A \quad (\text{B1})$$

As  $\alpha$  decreases income becomes less concentrated in the lower part of the distribution and it becomes more likely that consumers will be excluded from mainstream financial services. Alternatively, as the standard of living for the lowest income households,  $a$ , increases, the condition for exclusion is less likely to hold.

Cutoff income,  $w_p^*$ , and the level of exclusion,  $G(w_p^*)$ , under a Pareto distribution are given by:

$$w_p^* = \frac{c_B - c_A}{r + \theta \left(1 - \frac{1}{\alpha}\right)} \Rightarrow G(w_p^*) = 1 - \left( \frac{a \left[ r + \theta \left(1 - \frac{1}{\alpha}\right) \right]}{c_B - c_A} \right)^\alpha \quad (\text{B2})$$

We can show that the proportion of the population excluded,  $G(w_p^*)$ , is decreasing with the rate of return,  $r$ , and the income of the poorest consumer (our standard of living parameter),  $a$ . It is also straightforward to show that both the cutoff level of exclusion and the proportion of those excluded are decreasing with  $\alpha$ . The significance of  $\alpha$  as a measure of inequality is not clear. Pareto himself referred to  $\alpha$  as a measure of inequality. But if we measure inequality using the Gini coefficient, an increase in  $\alpha$  leads to a decrease in inequality (Chipman 1974).<sup>28</sup> An increase in  $\alpha$  represents an increase in the density at the lower tail of the distribution, but it also represents a fall in mean income.<sup>29</sup> In order to interpret the impact of changes in the Pareto distribution on exclusion in our model we need to consider both the standard of living parameter  $a$  and the shape of the curve,  $\alpha$ .

In the case of the US,  $\alpha$  has decreased over the last 30 years, leading to an increase in overall mean income. But as Atkinson, Piketty, and Saez (2011) argue, this rise in mean real income has been driven mainly by an increase in the right tail of the income distribution, while the standard of living of the lowest income households,  $a$ , has remained mostly unchanged.<sup>30</sup> This suggests the opposite of how Pareto interprets  $\alpha$ , meaning that a lower  $\alpha$  can be associated with greater

<sup>27</sup> There is no upper condition as our income distribution does not have a finite upper limit.

<sup>28</sup> Under a Pareto distribution, the Gini coefficient is given by:  $G = \frac{1}{2\alpha - 1}$ .

<sup>29</sup> Mean income under a Pareto distribution is equal to:  $a \left( \frac{\alpha}{\alpha - 1} \right)$ .

<sup>30</sup> Atkinson, Piketty, and Saez (2011) show that although over the previous 30 years real income had grown at an average annual rate of 1.2%, the majority of that growth had been due to the growth in income of the top 1% of the population.

inequality.<sup>31</sup> On the other hand, their study of real income in the UK found that although  $\alpha$  has been decreasing, the standard of living for the lowest-income households,  $a$ , has increased.<sup>32</sup>

Based on these results our model would predict that in the US exclusion from mainstream banking must have increased over the last 30 years. From (B2) we can see that, holding everything else constant, decreasing  $\alpha$  without an increase in  $a$  would lead to greater exclusion. In the case of the UK the prediction of the model would be ambiguous. As we argued above, a decrease in  $\alpha$  would lead to greater exclusion, while an increase in  $a$  would cause exclusion to decrease. Interestingly, the proportion of the unbanked in the UK decreased steadily from 2002 to 2011 (Rowlingson and McKay 2017). This trend might suggest that in the UK the impact of a rise in  $a$  has outweighed a fall in  $\alpha$ .<sup>33</sup>

## Appendix C: Results from structural estimation

---

**31** In fact, Atkinson, Piketty, and Saez (2011) demonstrate that when using top-income data the inverse of  $\alpha$  is a measure of inequality.

**32** Their study found similar results for most English speaking countries, and to a smaller extent for some Nordic countries.

**33** Note that these results might also be due to a variety of other factors, such as changes in the rate of return available to banks,  $r$ , as well as efforts by the UK government to increase access to banking.

Table C1: Parameter estimates.

	POOL	RE1	FE1	RE2	FE2
$\theta$	0.0091*** (0.0003) [0.0083,0.0100]	0.0094*** (0.0004) [0.0085,0.0103]	0.0090*** (0.0003) [0.0082,0.0098]	0.0091*** (0.0003) [0.0083,0.0101]	0.0090*** (0.0003) [0.0082,0.0098]
$\phi_1$		0.2037*** (0.0302) [0.1457,0.2936]		0.1830*** (0.0279) [0.1274,0.2636]	
$\phi_2$				0.0423*** (0.0085) [0.0264,0.0761]	
$\sigma_a^2 \times 1000$		0.0117*** (0.0016) [0.0084,0.0171]		0.0118*** (0.0016) [0.0086,0.0170]	
$\sigma_\lambda^2 \times 1000$				0.0010*** (0.0003) [0.0005,0.0017]	
$\alpha$ fixed effects	No	No	Yes	No	Yes
$\lambda$ fixed effects	No	No	No	No	Yes
Log likelihood	-9473.66	-9336.07	-9003.32	-9301.93	-8947.19
LR test, POOL		275.18***	940.68***	343.45***	1052.93***
LR test, RE1			665.49***	68.27***	
LR test, FE1					112.26***
LR test, RE2					709.48***

The column POOL correspond to pooled data, the column RE1 to one-way random effects, the column FE1 to one-way fixed effects, the column RE2 to two-way random effects and the column FE2 to two-way fixed effects. The number of observations is 976 in all five estimations. Clustered bootstrapped standard errors are given in parenthesis and bias corrected and accelerated clustered bootstrap confidence intervals at the 1% significance level are given in square brackets. LR test stands for likelihood ratio test vis-à-vis the relevant specification. \*\*\* indicates coefficients significantly different from zero at 1% level.

## References

- Amaral, P. S. 2017. *Monetary Policy and Inequality*. Federal Reserve Bank of Cleveland: Economic Commentary.
- Ampudia, M., D. Georgarakos, J. Slacalek, O. Tristani, P. Vermeulen, and G. L. Violante. 2018. "Monetary Policy and Household Inequality." In *Working Paper Series 2170*. Frankfurt: European Central Bank.
- Atkinson, A. B. 1995. "Capabilities, Exclusion, and the Supply of Goods." In *Choice, Welfare, and Development*, edited by K. Basu, P. K. Pattanaik, and K. Suzumura. chapter 2, 17–31. Oxford: Oxford University Press.
- Atkinson, A. B., T. Piketty, and E. Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49 (1): 3–71.
- Bakker, T., N. Kelly, J. Leary, and E. Naypal. 2014. "Data Point: Checking Account Overdraft." In *Technical report*. Washington, D.C.: Consumer Financial Protection Bureau (CFPB).
- Basu, S., and J. C. Wang. 2007. *Technological Progress, "Money" in the Utility Function, and the "User Cost of Money"*. Boston: mimeo.
- Beck, T., A. Demircuc-Kunt, and M. S. M. Peria. 2007. "Reaching Out: Access to and Use of Banking Services across Countries." *Journal of Financial Economics* 85: 234–66.
- Beck, T., A. Demircuc-Kunt, and M. S. M. Peria. 2008. "Banking Services for Everyone? Barriers to Bank Access and Use Around the World." *The World Bank Economic Review* 22 (3): 397–430.
- Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.
- Bertrand, M., and A. Morse. 2011. "Information Disclosure, Cognitive Biases, and Payday Borrowing." *The Journal of Finance* 66 (6): 1865–93.
- Bord, V. M. 2017. *Bank Consolidation and Financial Inclusion: The Adverse Effects of Bank Mergers on Depositors*. Cambridge, Massachusetts: Harvard University.
- Carrell, S., and J. Zinman. 2014. "In Harms Way? Payday Loan Access and Military Personnel Performance." *Review of Financial Studies* 27 (9): 2805–40.
- Celierier, C., and A. Matray. 2019. "Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation." *The Review of Financial Studies* 32 (12): 4767–809.
- Chipman, J. S. 1974. "The Welfare Ranking of Pareto Distributions." *Journal of Economic Theory* 9 (3): 275–82.
- Dash, E. 2011. *Profits Falling, Banks Confront a Leaner Future*, A1. New York: The New York Times.
- Demircuc-Kunt, A., L. Klapper, D. Singer, and P. V. Oudheusden. 2015. "The Global Findex Database 2014, Measuring Financial Inclusion Around the World." In *Technical report*. Washington, D.C.: World Bank Development Research Group.
- Dick, A. A. 2008. "Demand Estimation and Consumer Welfare in the Banking Industry." *Journal of Banking & Finance* 32 (8): 1661–76.
- FDIC. 2015. "2015 FDIC National Survey of Unbanked and Underbanked Households." In *Technical report*. Washington, D.C.: Federal Deposit Insurance Corporation.
- FDIC. 2017. "2017 FDIC National Survey of Unbanked and Underbanked Households." In *Technical report*. Washington, D.C.: Federal Deposit Insurance Corporation.
- Flannery, M., and K. Samolyk. 2005. "Payday Lending: Do the Costs Justify the Price?" In *Working Paper 2005-09*. Washington, D.C.: FDIC Center for Financial Research.

- Freixas, X. 2008. *Microeconomics of Banking*, 2nd ed. Cambridge, Massachusetts: The MIT Press.
- FSA. 2000. "In or Out? Financial Exclusion: A Literature and Research Review." In *Consumer Research Paper 3*. London: Financial Services Authority.
- Gabszewicz, J. J., and J. F. Thisse. 1979. "Price Competition, Quality and Income Disparities." *Journal of Economic Theory* 20 (3): 340–59.
- Haberfeld, R. 2002. "Cognitive Dissonance, Microeconomics, and Checking Accounts." Also available at [http://chexsys.tripod.com/\\_article\\_bankstocks.html](http://chexsys.tripod.com/_article_bankstocks.html).
- Ho, K., and J. Ishii. 2011. "Location and Competition in Retail Banking." *International Journal of Industrial Organization* 29 (5): 537–46.
- Jackson, G. 2019. *Londoners take to Cashless Payments as North-South Gap Emerges*. London: The Financial Times.
- Kempson, E., and C. Whyley. 1998. *Access to Current Accounts*. London: British Bankers Association.
- Lenza, M., and J. Slacalek. 2021. "How Does Monetary Policy Affect Income and Wealth Inequality? Evidence from Quantitative Easing in the Euro Area." In *CEPR Discussion Papers 16079, C.E.P.R. Discussion Papers*.
- O'Farrell, R., and L. Rawdanowicz. 2017. "Monetary Policy and Inequality: Financial Channels." *International Finance* 20 (2): 174–88.
- Parekh, A., T. MacInnes, and P. Kenway. 2010. "Monitoring Poverty and Social Exclusion." In *Technical report*. York: Joseph Rowntree Foundation.
- Rowlingson, K., and S. McKay. 2017. "Financial Inclusion Annual Monitoring Report." In *Technical report*. Birmingham: University of Birmingham and Friends Provident Foundation.
- Shaked, A., and J. Sutton. 1982. "Relaxing Price Competition through Product Differentiation." *The Review of Economic Studies* 49 (1): 3–13.
- Skiba, P., and J. Tobacman. 2007. *The Profitability of Payday Loans*. Nashville, TN: Vanderbilt University.
- Son, H., and M. Tighe. 2011. *Bofa's Moynihan Says "new Normal" in Banking Not as Profitable*. New York: Bloomberg News.
- Stango, V., and J. Zinman. 2014. "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees." *Review of Financial Studies* 27 (4): 990–1030.