Università Commerciale "Luigi Bocconi"
PhD School

PhD program in Economics and Finance
Cycle: 31st
Disciplinary Field (code): SECS-P/01

Three Essays in Banking and Labor Markets

Advisor: Basile Grassi

PhD Thesis by
Andrea Pasqualini
ID number: 3006510

Academic Year: 2019/2020

To my father Franco.

You are now gone, but your belief in me made my journey possible.
Your Respect, Honesty, Humbleness and Love will always inspire and drive me.

Acknowledgements

I am deeply indebted to Basile Grassi, my main advisor, for his continued advice and support. I also thank Filippo De Marco, Nicolas Serrano-Velarde and Giacomo Calzolari for valuable discussions and suggestions that contributed to the first essay in this thesis. Antonella Trigari, Diego Comin and Riccardo Franceschin are co-authors in the second and third essays.

I thank Bocconi University and the PhD School, for their continued financial support.

Thanks to all PhD students and graduates I met at Bocconi University during the last five years. The moments spent together, both professionally and recreationally, made for an unforgettable experience. Special mention goes to Jean-Gabriel Lauzier, Zijian Wang and Francesco Giovanardi.

I thank Melika, my valuable partner. Her charisma, entrepreneurial spirit and loving attention helped me get past the challenges the PhD program and life put me through.

To our dog Minus, for sitting on my laptop while I work and providing unconditional love.

I thank my dear friends Matteo, Kata and Tatiana, for their continued friendship.

Last, but certainly not least, I thank my family. Their unconditional love, care and support provide the best environment for growing up and proceeding in life.

Abstract

This thesis spans two fields: banking and labor markets. The first essay contributes to the former field, while the second and third essays contribute to the latter. The research questions that are broadly relevant to Macroeconomics. The second and third essays are based on joint work with Diego Comin, Riccardo Franceschin and Antonella Trigari.

In the first essay, I ask: how can we measure and disentangle market power on lending and deposit markets? What are the implications on the relationship between market power and financial stability? I revisit this old question by developing a new structural approach to the joint estimation of markups on lending rates and markdowns on deposit rates for all US depository institutions between 1992 and 2019. Markups (markdowns) are wedges between the observed price for the output (input) good and the price that would realize if the bank was a price taker on that market. Markups have been trending downwards over time, while markdowns have been increasing after the great recession and decreasing as recovery began. Bigger banks tend to exert more market power on lending markets, while smaller banks exert more power on deposit markets. However, markdowns play a larger role in the profitability of banks relative to markups. I show that Herfindahl-Hirschman Indices are positively correlated with markdowns on deposit rates, but negatively with markups on lending rates. I compute the Z-score and the O-score as measures of financial stability within US banks. I show that higher markups are associated with a lower bankruptcy probability. Instead, markdowns correlate positively with default probabilities. When considered jointly, markups and markdowns both correlate negatively with the probability of bankruptcy. These results show that the sources of market power are important in addressing this old question in the literature.

In the second essay, we seek to explain the differences in the time series of unemployment we observe across Germany, France, Spain and Italy. We write a standard Diamond-Mortensen-Pissarides (DMP) labor market model with search and matching frictions and we use it to assess differences in labor markets across Germany, France, Spain and Italy. We simulate the model feeding in exogenous shocks to aggregate productivity and to the discount factor. We obtain three main results. We first confirm the finding in Hall (2017) that financial returns are correlated to unemployment with European data, possibly more than labor productivity. Second, we find that discount factors are a promising source of variation to explain fluctuations in unemployment. Finally, we observe that the extent to which the DMP model explains the four countries' variations of unemployment depends on labor market institutions, as captured by the calibration, and on the degree of wage rigidity. However, the timing of fluctuations of unemployment is different across countries. This cannot be explained by different discounts, as they happen to be somewhat contemporaneous across countries.

Based on the results of the second essay, we attempt to characterize differences across labor market outcomes in Europe in the third essay. We do so by writing a labor market model where Fixed-Term Contracts (FTCs) and Open-Ended Contracts (OECs) can simultaneously arise in equilibrium. We write a labor market model à la Diamond-Mortensen-Pissarides that allows for heterogeneous match-specific productivities. A random productivity is drawn when a firm

and a worker meet, but the parties only observe a noisy signal initially. As long as the match persists, parties may perfectly learn the productivity with a Calvo-style lottery. At the match and based on the information they observe, workers and productivities may decide to reject the match, sign a FTC or an OEC.

Contents

Chapter 1

Markups, Markdowns and Financial Stability in the Banking Sector

1.1 Introduction

The great recession highlighted the importance of financial stability in the banking system. Academics and policy makers focused on assessing and improving financial stability. Many papers focused on the relationship between financial stability and market power banks have. Particularly on this topic, academics have not reached clear consensus.

On the other hand, recent literature in Economics has brought the importance of market power front and center, especially in Macroeconomics. The literature acknowledges market power on output markets and recently started to investigate market power on input markets. Firms may behave as monopsonistic firms, which lowers the observed price for their input goods relative to the perfect competition benchmark. In the banking literature, competition plays many roles and there is little agreement about the desirable state. Existing contributions highlight the importance of competition both on lending and deposit markets.

In this paper, I tackle both issues with a novel approach. I use the production approach to the estimation of markups and markdowns found in De Loecker and Warzynski (2012) and Morlacco (2019) to banking data. Given balance sheet and income statement data, the production approach allows for identification of markups and markdowns. The strength of this methodology lies in the data requirement. While markups are typically thought of as arising from the elasticity of demand for the output good and markdowns from the elasticity of supply for the input good, the production approach does not require demand and supply data. Originally based on Hall (1988), the production approach consists of a structural model of firm behavior. The first-order condition links the unobservable markup to expenditure shares, which can be observed directly, and production elasticities, which can be estimated. Morlacco (2019) expands this methodology in order to allow for buyer's market power, therefore introducing markdowns. Throughout this text, the markup is the wedge between the price for the output good one observes and the price one would observe were the seller a price taker. Conversely, the markdown is the wedge between the price for the input good we observe and the price that would realize if the buyer was a price taker. I jointly estimate both wedges for lending markets and deposit markets. For identification, I assume that banks hire labor on perfectly competitive markets and that both labor and deposits are not subject to adjustment costs.[1]

[1]The assumptions of no adjustment costs are credible given that I use yearly balance sheet data. At the yearly frequency, banks are unlikely to face significant frictions in hiring labor or taking deposits. On the other hand, while the first assumption goes against several papers documenting buyers' power on labor markets, it may be reasonable to believe that banks do not require highly specialized workers, such as tellers and administrative staff. They can

I find that markups on lending rates are trending downwards between 1992 and 2019, while markdowns on deposit rates have substantially increased after the great recession. The yearly cross-sectional dispersion of markups remains somewhat stable over the sample period, while the dispersion of markdowns follows the level: increases in the average (or median) markup are associated with increases in the cross-sectional dispersion, and vice versa. I compute the correlations with certain observable bank characteristics. Bigger banks tend to charge more markups and less markdowns. The same holds for more leveraged banks. Profitability correlates positively with both markups and markdowns. Banks that have a higher share of loans among their earning assets charge higher markups, but lower markdowns. Finally, banks that pay a larger share of their income as dividends are associated with higher markups and markdowns.

I compare markups and markdowns to the Herfindahl-Hirschman Index (HHI) computed at the state level for both deposits and loans. HHIs are usually taken as measures of market power, both for banks and manufacturing firms. I show that the HHI on deposits correlates positively with markdowns on deposit rates, although weakly. Instead, the HHI on loans correlates negatively with markups on lending rates. This suggests that markups and markdowns do not necessarily capture the same phenomena as HHIs.[2]

I obtain measures of financial stability by computing the Z-score (Altman, 1968) and the O-score (Ohlson, 1980). These consists of the predicted values of two regressions, where the left-hand side is an indicator of bankruptcy, which equals one if the bank will default within the following year. The two scores differ for the right-hand side term, which are generally balance sheet ratios. The scores can be interpreted as probabilities of bankruptcy that can be predicted (in sample) by reports of condition, such as balance sheets. I regress these measures of financial stability on the markups and markdowns I obtained. The goal of such regression is to assess the *correlation* between market power and financial stability. Interestingly, I find that markups correlate negatively with default probabilities. The correlation is small, although significant. I also find that markdowns are positively correlated with default probabilities, although the sign reverses once I account for markups as well. While I control for bank fixed effects, this result may be driven by a generalized increase in bank size that occurs across the whole cross section. Bigger banks rely more on markups than markdowns and, at the same time, there has been considerable mergers or acquisitions activity throughout the sample period.

This paper contributes to two main strands of the literature. The first one deals with the estimation or computation of measures of competition. The production approach to the estimation of markups was first introduced by Hall (1988). De Loecker and Warzynski (2012) later developed the full procedure, which relies on the estimation of the production function. More recently, De Loecker and Eeckhout (2019) and related papers from the same authors employed the methodology to document global trend in markups across various industries. I contribute to this literature by providing evidence that more narrowly focuses on banks. This is motivated by the fact that the concept of production function does not obviously relate to banks. Consequently, I adapt the estimation of production functions to suit banking data. Morlacco (2019) modified the technique in De Loecker and Warzynski (2012) in order to accommodate market power in input markets. I adapt their techniques to banking data in order to estimate the markdown on deposit rates. While the aforementioned papers deal with manufacturing firms, this paper brings the same insights to the banking sector.

The second strand of the literature I contribute to deals with the relationship between financial stability and market power. This literature contains conflicting results, both theoretically

therefore hire labor from a pool where banks compete with other industries.

[2]Future revisions of this paper will include a comparison with the Boone (2008) indicator. This alternative measure of competition relies on the relationship between profits and marginal costs. For a given increase in marginal costs, profits should decrease. The bigger the decrease, and the harsher competition there is in the market. The argument supporting this hypothesis is due to firm efficiency.

and empirically. For example, Hellmann et al. (2000) argue that competition may be detrimental for financial stability. In a scarcely competitive environment, banks realize profits, which can be accumulated and may serve as buffer against adverse shocks. Conversely, Boyd and De Nicoló (2005) argue that competition may foster financial stability. Banks with more market power do realize more profits, but also induce higher loan interest rates. This may increase risk-taking attitudes of firms that apply for loans. Empirically, Beck et al. (2006) show that more concentrated banking systems are associated with more financially stable economies. Instead, Schaeck et al. (2009) use the H-statistic of Panzar and Rosse (1987) to find that more competitive environments are associated with more stable banks. Berger et al. (2009) find that banks with more market power have less risk exposure. While this literature tends to focus on lending markets, I disentangle the effect of market power on lending and deposit markets. I *jointly* estimate markups and markdowns. I show that higher markups are associated with lower bank default probabilities, while the opposite holds for markdowns. Once I account for both markups and markdowns, I find that they both correlate negatively with the probability of bankruptcy. This change of sign for markdowns provides further motivation to the approach: it is important to disentangle market power for output goods and market power for inputs. Any alternative method that does not disentangle these two sources of market power may misinterpret the role of competition for financial stability.

The rest of the paper is organized as follows. Section 1.2 details the production approach to the estimation of markups and markdowns, together with its application to the banking sector. Section 1.3 presents the data I use. Section 1.4 shows the empirical results on markups on lending rates and markdowns on deposit rates, together with correlations with observable bank characteristics and a comparison with the Herfindahl-Hirschman Indices. Section 1.5 shows details on the measures of financial stability I use. Section 1.6 presents the results that relate market power on financial stability. Section 1.7 concludes.

1.2 The production approach to markups and markdowns estimation

The production approach to the estimation of markups and markdowns is detailed in De Loecker and Warzynski (2012) and Morlacco (2019). It relies on a simple structural model of firm behavior given a cost function and a production function. Before I detail the methodology, it is useful to clarify its use in the context of banking.

1.2.1 Conceptual framework

Production functions typically belong to the realm of manufacturing firms. They describe the transformation from inputs goods into output goods in a concise way. They are often referred to as "black boxes" because they do not describe how exactly such transformation takes place. Production functions are rarely encountered in the literature on banking: existing papers have focused on the role of banks as intermediaries, exploring the economic mechanisms that justify such roles. I do not model explicitly any specific role of banks. Instead, I take those roles for granted and I model them with a production function. This function for banks includes all the economic frictions that a bank addresses, such as informational asymmetries. In this sense a loan production function can be seen as a reduced-form characterization of the activity of the bank that does not ignore their economic role.

I assume that banks collect sources of financing, such as deposits and equity, and use them to provide loans. To do so, banks also need traditional input goods, such as labor and capital. This approach follows directly Sealey and Lindley (1977), who characterize the activity of a bank in terms of classical production theory. In their paper, the authors also describe the main difference between a bank and a manufacturing firm in terms of production. A manufacturing firm requires capital and labor to produce a physical good. Instead, banks *also* require sources of financing in order to supply a loan. To see this, consider a bank in a frictionless, simplified world that only uses deposits and labor and that such inputs are already being efficiently exhausted. Suppose that this bank can hire more labor, but cannot raise an additional unit of deposits. In this case, the bank cannot increase its outstanding loans, because the balance sheet constraint binds. On the other hand, if such bank can source additional deposits but cannot hire more labor, then it does not have the capacity to process more loans. Therefore, there is a degree of complementarity between sources of financing and physical input goods. Generally speaking, the production feasibility set of a bank is affected by the balance sheet constraint, the need for physical goods (e.g., premises, IT equipment, labor), regulatory constraints and the sources of risk, such as a creditor's default risk and the risk of bank runs.

In reality, banks also use sources of financing other than deposits, such as equity and, in some cases, corporate bonds. Some banks can also be seen as multi-output firms, because some buy financial assets and repackage them as securities to be either held or traded. Because I focus on depository institutions, most of earning assets in those banks are made of loans, as the summary statistics below show. Additionally, equity and deposits almost entirely describe the liabilities side of balance sheets. For these reasons, I focus on a single-product production function, where the output is loans, and I restrict my attention to deposits and equity as the only sources of finances.

1.2.2 The structural model

Building on De Loecker and Warzynski (2012) and Morlacco (2019), I assume that every bank i in period t solves the following static cost minimization problem, subject to a production function and for a given level of outstanding loans:

$$\begin{aligned} \min_{D_{it},E_{it},N_{it},K_{it}} \quad & r^D_{it} D_{it} + r^E_{it} E_{it} + w_{it} N_{it} + r^K_{it} K_{it} \\ \text{subject to} \quad & L_{it} = F\left(D_{it}, E_{it}, N_{it}, K_{it}\right), \end{aligned} \tag{1.1}$$

where L_{it} are loans, D_{it} are deposits, E_{it} is equity, N_{it} is labor, K_{it} is capital, r^D_{it}, r^E_{it}, and r^K_{it} are the input interest rates paid on deposits, equity and capital respectively, w_{it} is the wage per efficient labor unit and $F(\cdot)$ is the loan production function. Given a level of outstanding loans L_{it}, the solution to this problem characterizes the optimal mix of physical input goods and financial assets to use in loan production. In order to identify the markup on lending rates, it is necessary to assume price-taking behavior and no adjustment costs for at least one input good or asset. Conversely, in order to identify the markdown on deposit rates, it is necessary to have an identified measure of markup on lending rates and, additionally, to assume that banks are not price-takers on deposit markets and that deposits are not subject to adjustment costs.

I assume that labor satisfies the required assumptions for identification of markups on lending rates. Two reasons justify this assumption. First, it is arguably the case that banks are not price-takers on the markets for deposits, while equity and capital may well be subject to adjustment costs. In particular, one objective of this paper is to identify a markdown on deposit rates. Second, banks compete with other industries for low-skilled workers and for administrative staff on one hand. On the other, banks may not perfectly compete with other industries for mid- or top-management workers. However, the more one climbs the job ladder within a bank and the more she is likely to be also paid with other forms of compensation, such as stock options, rather than wages.

The first order condition for (1.1) with respect to labor is

$$w_{it} = \lambda_{it}\frac{\partial F}{\partial N_{it}},$$

where λ_{it} is the Lagrange multiplier associated to the production function and corresponds to the marginal cost of loan production. By multiplying each side by $N_{it}/(r^L_{it}L_{it})$, where r^L_{it} is the interest rate on loans a bank charges, and rearranging terms we obtain the following expression:

$$\underbrace{\frac{r^L_{it}}{\lambda_{it}}}_{\mu_{it}} = \underbrace{\left[\frac{\partial F}{\partial N_{it}} \cdot \frac{N_{it}}{L_{it}}\right]}_{\theta^N_{it}} \cdot \underbrace{\left[\frac{w_{it}N_{it}}{r^L_{it}L_{it}}\right]^{-1}}_{1/\alpha^N_{it}}. \tag{1.2}$$

The left-hand side is the interest rate on loans divided by the marginal cost of loan production, that is the gross markup on the lending rate, μ_{it}. The right-hand side is made of two components: the first, θ^N_{it}, is the elasticity of loan production to labor and the second, $1/\alpha^N_{it}$, is the inverse expenditure share of labor relative to loan interest income. This expression has operational content. It implies that we can compute the unobservable markup given the inverse expenditure share, which is readily available in balance sheet data, and the production function elasticity, which can be estimated given a panel of banks.

Consider now the first-order condition to (1.1) with respect to deposits D_{it}. Repeating the steps taken above, the first-order condition is

$$\underbrace{\left[1 + \frac{\partial r^D_{it}}{\partial D_{it}} \cdot \frac{D_{it}}{r^D_{it}}\right]}_{\psi_{it}} \cdot \underbrace{\frac{r^L_{it}}{\lambda_{it}}}_{\mu_{it}} = \underbrace{\left[\frac{\partial F}{\partial D_{it}} \cdot \frac{D_{it}}{L_{it}}\right]}_{\theta^D_{it}} \cdot \underbrace{\left[\frac{r^D_{it}D_{it}}{r^L_{it}L_{it}}\right]^{-1}}_{1/\alpha^D_{it}}, \tag{1.3}$$

where the term ψ_{it} relates to the inverse supply elasticity of deposits. Morlacco (2019), who focuses on French manufacturing firms, interprets ψ_{it} as a *markdown*. Formally, the markdown is here defined as the wedge between the deposit rate banks pay relative to the interest rate that banks would pay if they were price-takers on deposit markets. In order to empirically recover the markdown component, one needs the same ingredients as before (i.e., the inverse expenditure share and the elasticity of the production function) and, additionally, a measure of markups. This happens because the markdown component is unobservable given balance sheet data. Assuming that Equation (1.2) identifies the markup, one can take the ratio between Equation (1.3) and (1.2) and identify the markdown:

$$\frac{\theta^D_{it}/\alpha^D_{it}}{\theta^N_{it}/\alpha^N_{it}} = \frac{\mu_{it}\psi_{it}}{\mu_{it}} = \psi_{it}. \tag{1.4}$$

To recap, the production approach to the estimation of markups and markdowns consists of three steps. The first is estimating the production function elasticity with respect to each input good. This exercise is standard in the Empirical Industrial Organization literature. The second is computing the inverse expenditure shares of each input good, which is trivial given income statement variables. The third and final step is choosing two input goods such that one is subject to neither monopsonistic competition nor adjustment costs and the other is not subject to adjustment costs. The first-order condition with respect to the first good allows for identification of the markup on lending rates, while the ratio of first-order conditions identifies the markdown. While the second and third steps are trivial, the first one requires some attention and I therefore turn to it now.

1.2.3 Estimation of the production function and production elasticities

Consider the following net loan production function for bank i at period t, with variables written in logs:

$$l_{it} = f(d_{it}, e_{it}, n_{it}, k_{it}; \beta) + \tilde{\varepsilon}_{it}, \tag{1.5}$$

where β is the vector of production function parameters and $\tilde{\varepsilon}_{it}$ is normally referred to as an unobserved productivity term. In the context of financial assets as loans, it is not clear what productivity means. In this paper, I assume that net loans are subject to repayment shocks. When $\tilde{\varepsilon}_{it}$ is positive, the bank receives a repayment from a loan that was not expected to realize. Conversely, when $\tilde{\varepsilon}_{it}$ is negative, the bank does not receive a repayment from a loan that was instead expected to realize. Because each bank has private information regarding the repayment shocks $\tilde{\varepsilon}_{it}$ while the researcher does not, the OLS estimate of β will be subject to an endogeneity issue because of an omitted variable bias. Therefore, the OLS estimate of β cannot be reliably interpreted as (the vector of) the loan elasticities to the input goods. This endogeneity issue is well documented in Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015).

In this paper, I follow the approach of Ackerberg et al. (2015) in estimating the production function parameters. Suppose that the repayment shock $\tilde{\varepsilon}_{it}$ can be decomposed in two additive terms (in logs): a term that is known by bank i at time t, ω_{it} and a term that is unknown to both the bank and the researcher, ε_{it}. Therefore, we can rewrite Equation (1.5) as

$$l_{it} = f(d_{it}, e_{it}, n_{it}, k_{it}; \beta) + \omega_{it} + \varepsilon_{it}.$$

To identify the production function parameters β, Ackerberg et al. (2015) propose a two-step GMM approach. Identification of β occurs at the second stage. Suppose that there exists an intermediate, complementary production good or asset that the bank chooses based also on the privately observed term ω_{it}. Such term does not appear in the production function (1.5) because this is a value-added specification. In the context of manufacturing firms, such good can be materials, and is therefore denoted as m_{it}. In the context of banks, there is no such thing as materials. However, I write net loans in the value-added production function. A control variable for banks that appears in the balance sheets is the loan loss provisions, which reflects the fraction of repayments of gross loans the banks deems noncollectable. The determination of each period's loan loss provision is up to each bank, as each one is expected to have private information about its customers. Let m_{it} be determined by the following demand function:

$$m_{it} = h(d_{it}, e_{it}, n_{it}, k_{it}, \omega_{it}). \tag{1.6}$$

This demand function arises from the optimization problem in (1.1) when the constraint is not the value-added production function but, rather, the gross production function, where intermediate goods or assets would appear. Assume that the function h is invertible with respect to ω_{it}, such that we can write

$$\omega_{it} = h^{-1}(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}).$$

Plug this expression in Equation (1.5) to obtain the following:

$$\begin{aligned} l_{it} &= f(d_{it}, e_{it}, n_{it}, k_{it}; \beta) + h^{-1}(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}) + \varepsilon_{it} \\ &= \Phi(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}; \beta) + \varepsilon_{it}. \end{aligned} \tag{1.7}$$

Equation (1.7) constitutes the first step in the estimation procedure. It is estimated with OLS where the function $\Phi(\cdot)$ is approximated with a n-th order polynomial. Let Φ_{it} denote the predicted values of the regression.

The second step consists of a GMM estimation. From Equation (1.5), and given a value for β, we have that

$$\omega_{it}(\beta) = \Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right).$$

A sufficient condition for identification of β is that ω_{it} follows a Markov process at the bank level. For concreteness, I assume that the repayment shock ω_{it} follows an AR(1) process for each bank i:

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it}, \tag{1.8}$$

where the innovation term ξ_{it} is not in the information set of each bank. The GMM moment condition requires that

$$\begin{aligned} 0 &= \mathbf{E}_{it}\left(\xi_{it}(\beta)\right) \\ 0 &= \mathbf{E}_{it}\left(\Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) - \rho\left[\Phi_{it-1} - f\left(d_{it-1}, e_{it-1}, n_{it-1}, k_{it-1}; \beta\right)\right]\right). \end{aligned}$$

Note that the expectation is conditional on the information set of bank i at time t, which makes it operationally difficult to deal with. Following Ackerberg et al. (2015), I instrument the conditioning information set with a vector of variables that I assume not to correlate with ξ_{it}. In particular, the instrumented GMM condition is

$$\mathbf{E}\left(\left[\Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) - \rho\left[\Phi_{it-1} - f\left(d_{it-1}, e_{it-1}, n_{it-1}, k_{it-1}; \beta\right)\right]\right] \otimes \begin{bmatrix} 1 \\ k_{it} \\ l_{it-1} \\ \Phi_{it-1} \end{bmatrix}\right) = 0. \tag{1.9}$$

Operationally, the variable l_{it} is the log of total net outstanding loans, d_{it} is the log of total domestic deposits, e_{it} is the log of total equity, n_{it} is the log-expenditure on labor and k_{it} is the log of premises and equipment. I specify the production function (1.5) to be Cobb-Douglas. I implement the second step in the estimation of the production function with a numerical root-finding routine. I set the initial condition for β to the OLS estimate of Equation (1.5).

Assuming a Cobb-Douglas production function implies that the loan elasticities to each input good or asset are constants across time and banks. Hence, the cross-sectional and time-series features of markups on lending rates are driven by the features of the expenditure shares. The elasticities simply rescale the expenditure shares. This is easily seen in Equation (1.2).

1.3 Data and summary statistics

I use data from the Federal Deposit Insurance Corporation (FDIC). They maintain and provide the Statistics on Depository Institutions (SDI). These are balance sheet, income statement and demographic variables available at the quarterly frequency. They are compiled from the quarterly Call Reports, which are reports of condition and income. Each depository institution is required by law to fill the Call Reports. The structure of the filled form and the amount of detail in reported information depends on the amount of total assets and on whether banks have only domestic or domestic and foreign offices. The data are publicly available starting from 1992Q4. I use figures from 1992Q4 to 2019Q4. Income statement variables are cumulated within each fiscal year for each bank. When I take within year-bank differences to obtain the non-cumulated version of the variables, I observe significant seasonal variation at the year-bank level for all income statement variables. For example, cash dividends are typically registered in the income statement only at

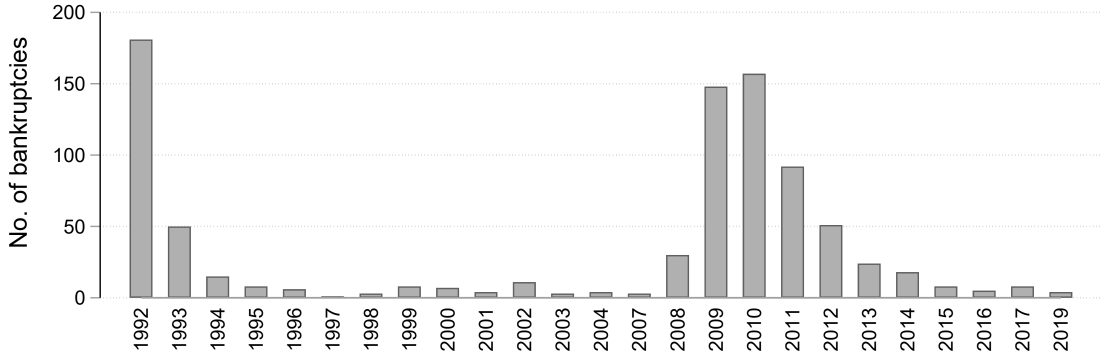

Figure 1.1: Bankruptcies in sample period.

the end of the year. For this reason, I focus my attention to end-year observations, effectively using data at the yearly frequency.

I use data from the Bank Failure and Assistance database, also provided by the FDIC, in order to obtain an indicator of bankruptcy for each bank. These data report the date of bankruptcy for every failed bank to date. Additionally, the FDIC reports the type of settlement after the bankruptcy: failure or assistance. In the former case, the financial institution is liquidated. In the latter, the FDIC provides guidance so that the bankrupt institution is acquired by another bank. From these data, I compile an indicator variable for each bank present in the SDI. The variable equals one if, in that year, the bank went bankrupt.

Table 1.1 shows summary statistics for the cross-sections of banks in 1992 and 2019. Their comparison provides an indication of changes in the US banking industry over 27 years. Each table groups statistics by percentile brackets of total assets. The number of depository institutions in the US went from 13,973 in 1992 to 5,186 in 2019. Roughly 48 percent of total banking system assets belonged to the top size percentile in 1992 and roughly 75 percent in 2019. The number of existing banks reduced over time. This is a first indicator that banking activities concentrated in bigger banks over time.

The composition of the balance sheets of banks has slightly changed over time. Net loans represented 51 to 58 percent of total assets in 1992 and 59 to 70 percent in 2019. Compared to held securities, net loans make for the majority of earning assets. Financing is primarily given by deposits, which backed roughly 73 to 88 percent of total assets in 1992 and 77 to 83 percent in 2019. Equity has become more important over time, backing roughly 7 to 9 percent of total assets in 1992 and 11 to 13 percent in 2019.

The primary source of income has always been interest from loans, ranging from roughly 55 to 64 percent of total income in 1992 and from 56 to 72 in 2019. Interest income from held securities decreased over time, from approximately 17 to 26 percent of total income in 1992 to 9 to 13 in 2019. Somewhat surprisingly, service charges on deposits are a relatively small source of income for banks, accounting for roughly 2-3 percent of total income both at the beginning and at the end of the sample. This reinforces the assumption that deposits are input goods. Although banks may compete over commission fees, they are not an important source of revenue for banks.

Figure 1.1 shows the number of bankruptcies in the sample period. There has been a number of bankruptcies at the beginning of the sample and right after the great recession. Of the 849 bankruptcies in the sample period, 640 resolved with other banks acquiring all deposits (insured and uninsured) and some assets, 126 resolved with other banks acquiring only insured deposits, 51 resulted in complete payouts, 17 resulted in other institutions paying insured deposits (without acquiring them) and 15 resulted in assisted transactions, where the FDIC managed transactions

Table 1.1: Summary statistics at the beginning and at the end of the sample period. Dollar figures are adjusted for inflation and expressed in terms of 2019 US dollars. The term "percentile bracket" refers to the cross-sectional distribution of total assets within each year.

	1992						2019					
Percentile bracket	[0, 75)	[75, 90)	[90, 95)	[95, 98)	[98, 99)	[99, 100]	[0, 75)	[75, 90)	[90, 95)	[95, 98)	[98, 99)	[99, 100]
No. of banks	10,479	2,096	699	419	140	140	3,889	778	259	156	52	52
Average total assets (bln USD)	0.089	0.370	0.900	2.448	6.258	28.050	0.200	0.958	2.537	8.014	25.297	269.025
Median total assets (bln USD)	0.075	0.344	0.843	2.196	6.136	16.237	0.160	0.897	2.318	6.810	24.757	117.848
Average income (bln USD)	0.008	0.031	0.077	0.206	0.547	2.602	0.011	0.052	0.138	0.411	1.290	14.148
Median income (bln USD)	0.006	0.029	0.071	0.179	0.498	1.431	0.008	0.037	0.093	0.264	0.816	9.000
Average expense (bln USD)	0.006	0.025	0.062	0.169	0.429	2.091	0.008	0.045	0.115	0.341	1.271	6.492
Median expense (bln USD)	0.005	0.023	0.057	0.144	0.398	1.162	0.006	0.032	0.077	0.212	0.739	4.041
Average NIM / assets (%)	4.227	3.935	3.863	3.674	3.607	3.576	3.460	3.423	3.325	3.305	3.293	3.144
Percentage of total system assets (%)	11.386	9.503	7.710	12.566	10.732	48.102	4.153	3.977	3.507	6.673	7.021	74.669
Average percentage relative to total assets in size category												
Net loans	51.821	57.297	59.242	57.662	57.984	58.907	63.186	70.388	69.810	70.024	68.123	59.389
Securities	31.508	28.643	26.603	26.641	24.572	21.420	19.057	16.392	16.666	17.806	17.074	21.885
Intangible capital	0.127	0.193	0.320	0.602	0.668	0.604	1.601	1.705	1.476	1.261	1.215	0.797
Physical capital	1.603	1.588	1.396	1.318	1.264	1.286	0.330	0.836	1.113	2.069	3.378	2.094
Deposits	88.194	87.272	84.978	80.650	75.862	73.357	83.179	82.696	81.198	78.660	76.812	77.338
Equity	9.559	8.221	7.415	7.403	7.284	6.889	13.096	11.973	11.729	12.634	13.037	11.191
Average percentage relative to income in size category												
Int. income from loans	60.655	63.799	64.486	62.516	59.799	55.750	71.548	72.198	69.929	69.366	69.500	56.787
Int. income from securities	26.681	24.112	22.168	23.580	18.777	17.344	10.717	8.755	9.392	10.962	9.026	12.850
Int. income from lease financing receivables	0.151	0.195	0.402	0.488	1.224	1.322	0.194	0.213	0.370	0.437	0.500	1.029
Income from charges on deposits	3.017	2.557	2.373	2.441	2.912	2.797	3.607	3.150	3.348	3.152	3.427	3.158
Average percentage relative to expenses in size category												
Int. expense on deposits	51.196	52.147	49.648	45.851	40.652	34.539	19.053	21.197	22.631	23.381	21.623	22.646
Wages and salaries	23.090	20.742	19.612	18.063	18.493	19.750	44.995	44.264	41.817	38.733	35.294	32.269
Cash dividends	6.343	6.293	5.163	5.344	4.856	5.923	16.964	16.728	16.882	28.960	27.209	36.589

across banks such that the bankrupt institution's charter survives. These numbers highlight the relevance of acquisitions or assumptions relative to payouts or assisted transactions.

1.4 Markups and markdowns in the banking industry

In this section, I present the results on the estimation of markups and markdowns. I also describe their correlation with observable bank characteristics and I compare them with Herfindahl-Hirschman indices.[3]

1.4.1 Evolution of markups and markdowns over time

Following the procedure detailed in Section 1.2, I compute the markups and markdowns on lending and deposit rates respectively. Table 1.2 reports the estimates of the production function parameters, assuming that the production function is Cobb-Douglas. I report the results using OLS on Equation (1.5), together with the GMM results using the procedure described above.[4] The estimates also represent the production function elasticities with respect to each input good *ceteris paribus*, because of the functional form and because all variables are expressed in logs.

First, the OLS and GMM estimates are different. This comes from the fact that OLS estimates suffer an omitted variable bias. Second, the results show the predominant elasticity of loans with respect to deposits. A one percent increase in deposits translates into a roughly 0.54 percent increase in net loans. The elasticity of loans to equity is roughly 0.20, to labor is 0.28 and to physical capital is 0.03. For robustness, I also report the estimates using quarterly data. Quarterly income statement variables are cumulative at the bank-year level. For this reason, I take their first-differences within every bank-year pair. The results are qualitatively comparable, where the major differences occur for the elasticities of loans with respect to deposits and labor. The results on markups and markdowns that follow rely on the GMM estimates using the yearly data.

Figure 1.2 shows the time series behavior of the average, median and interquartile range of markups on lending rates. The average net markup went from roughly 50 percent in 1992 to roughly 25 percent in 2019. Overall, markups have been trending downwards over time, with a temporary increase before the great recession.

Figure 1.3 shows instead the time series of the average, median and interquartile range of markdowns on deposit rates. The average markdown in 2019 is roughly ten times higher than the average markdown in 1992. The increase predominantly occurred after the great recession. Importantly, the dispersion of markdowns has increased whenever the average increased, and vice versa.

As Equation (1.3) shows, there is a mechanical relationship between markdowns and markups. Keeping the expenditure share α_{it}^D and the elasticity θ_{it}^D constant, if the markup μ_{it} increases, then the markdown ψ_{it} has to decrease. However, the reported time series behavior suggest that there also is an economic interpretation to this relationship. We observe markups on lending rates to decrease with the great recession, while markdowns on deposit rates increase. We also observe that there has been an increase in the number of bankrupt banks during the great recession. Because policy makers were concerned with financial stability at the time, they had to impose more stringent rules about the risk-taking behavior of banks. This meant that banks faced harsher competition on lending rates and had to make up for the lost profitability in order to survive. Deposits have been a way for banks to sustain their profit streams, by paying deposit rates that are relatively lower than comparable interest rates (e.g., risk-free rates).

[3] I will add the comparison with the Boone indicator in future revisions of this work.

[4] The standard errors for the GMM estimates need to be bootstrapped and will appear in future revisions of this paper.

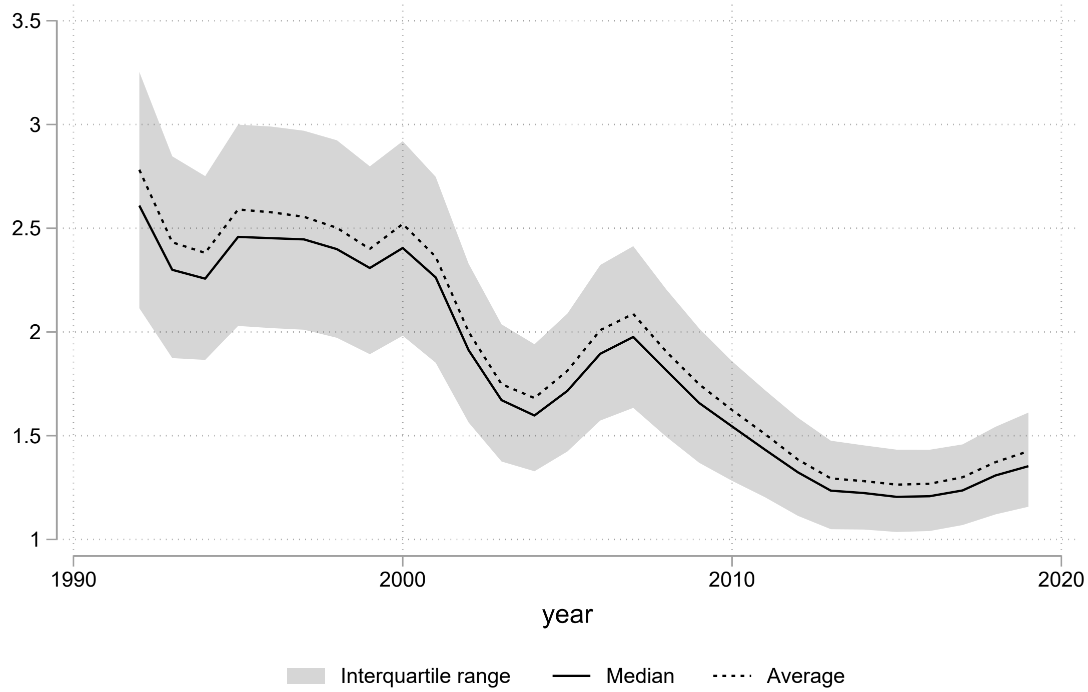

Figure 1.2: Average, median and interquartile range of gross markups across years. Each yearly cross-section of markups has been trimmed 1% top and 1% bottom to account for outliers.

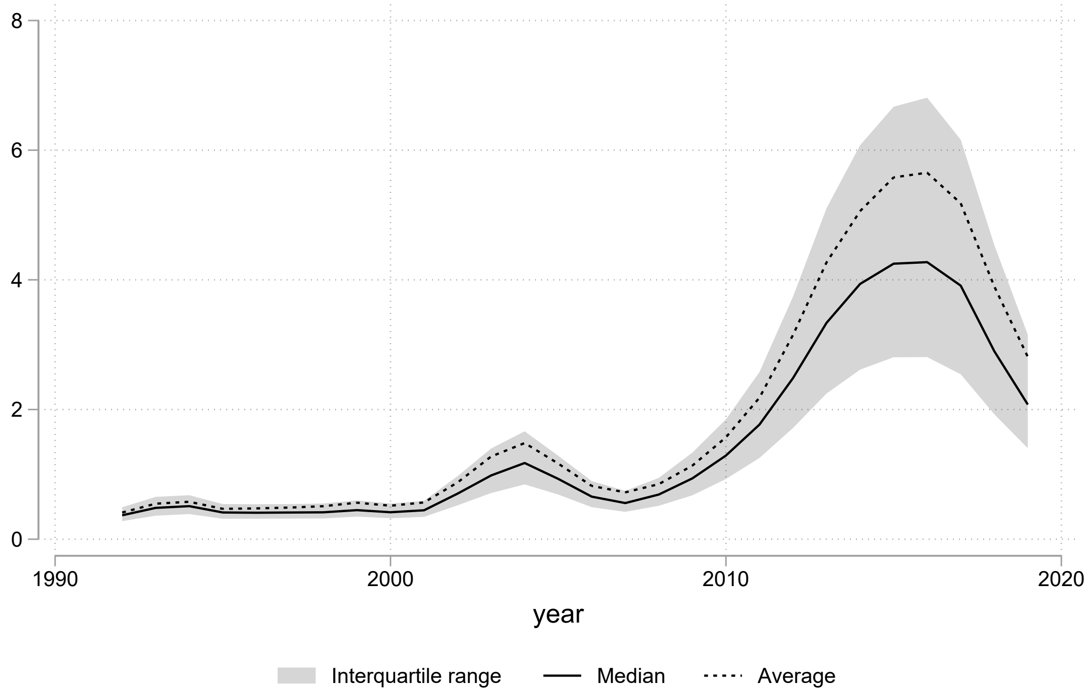

Figure 1.3: Average, median and interquartile range of gross markdowns across years. Each yearly cross-section of markdowns has been trimmed 1% top and 1% bottom to account for outliers.

Table 1.2: Estimates of the production function parameters. All variables are expressed in logs. The columns for the initial condition refer to the OLS estimates of the production function (1.5), which are then used as initial condition for the GMM optimization. The columns for GMM estimates show the final results. The number in parentheses are standard errors. The standard errors for the GMM estimates need to be bootstrapped and will appear in future revisions of this paper.

	Net loans			
	Yearly data		Quarterly data	
	OLS	GMM	OLS	GMM
Constant	0.1975	0.0707	0.0124	0.0072
	(0.009)		(0.005)	
Deposits	0.6376	0.5394	0.7028	0.6981
	(0.002)		(0.001)	
Equity	0.1812	0.2036	0.2120	0.2200
	(0.002)		(0.001)	
Labor	0.2082	0.2877	0.0847	0.0808
	(0.002)		(0.001)	
Capital	0.0451	0.0386	0.0631	0.0604
	(0.001)		(0.000)	

1.4.2 Markups, markdowns and bank characteristics

How do markups and markdowns correlate with observable bank characteristics? To answer this question, I regress them separately on a set of balance sheet and income statement variables. This exercise is useful to shed light on which banks are able to charge higher markups or markdowns. However, it does not help understand the determinants of market power in loan and deposit markets. In the right-hand side variables I include the log of assets to capture the size of the bank, the log of assets over equity to capture leverage, the log of the net interest margin over assets to capture bank profitability, the log of net loans over held securities to consider different ways banks use earning assets and the log of cash dividends over net income to capture the relationship between the bank and its stockholders. All variables are in logs in order to interpret the estimates as the percent change in the left-hand side variable associated with a one percent increase in the right-hand side variable. I consider all four combinations that result from including, or not, bank and/or year fixed effects. Including fixed effects gives a sense of variation across banks or periods, rather than within. For example, including time fixed effects controls for the effect a particular year may have on all observations (e.g., the great recession). Including bank fixed effects helps control for certain bank characteristics that are not captured by the regressors, such as geographical location.

Table 1.3 shows the results from a set of regressions where the dependent variable is either the log-markup or the log-markdown and the independent variables are a set of bank characteristics. Columns 1 to 4 relate to log-markups. The signs of the coefficients change depending on whether I consider bank and/or year fixed effects. The first column of the table considers neither bank nor year fixed effects. We see that increases in size, leverage, profitability are associated with increases in the markups. Conversely, markups correlate negatively with the prevalence of loans as earning assets and with the amount of income paid out as dividends. However, the adjusted R-squared is relatively low, at just above five percent. This means that the regressors I considered are not enough to capture all variation in markups within banks and within periods. Once I

Table 1.3: Correlations between log-markups, log-markdowns and observable bank characteristics. Markups and markdowns have been trimmed one percent both at the bottom and at the top of each yearly cross-section.

	Log(Markups)				Log(Markdowns)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(Assets)	0.0074***	0.0256***	−0.1272***	0.1258***	0.0860***	0.0229***	0.6201***	−0.2142***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.005)	(0.002)
Log(Assets / equity)	0.1839***	−0.0628***	0.2917***	0.0670***	−0.7124***	0.0571***	−0.9370***	−0.1760***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.008)	(0.005)	(0.011)	(0.005)
Log(NIM / assets)	0.2767***	−0.2320***	0.4498***	0.0412***	−0.6810***	0.8217***	−1.0631***	0.2455***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.011)	(0.006)	(0.013)	(0.005)
Log(Loans / securities)	−0.0541***	0.0258***	−0.0514***	0.0042***	0.1647***	−0.0673***	0.1365***	−0.0125***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
Log(Cash dividends / income)	−0.0153***	0.0329***	−0.0470***	0.0055***	0.1485***	0.0017	0.1820***	0.0109***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Constant	1.1023***	−0.1754***	2.9610***	−0.8327***	−1.5514***	2.2289***	−8.6127***	3.7014***
	(0.016)	(0.013)	(0.022)	(0.018)	(0.041)	(0.022)	(0.065)	(0.032)
Observations	159139	159139	158180	158180	158924	158924	157985	157985
Bank FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.0541	0.4563	0.6066	0.8302	0.1355	0.7668	0.5092	0.9235

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

consider time fixed effects, all signs are flipped, except for the one with respect to bank size. The adjusted R-squared jumps up to roughly 45 percent. The regressors do a fair job at capturing some variation in markups across periods, but within banks. When I allow for bank fixed effects only, we observe that markups are negatively correlated with bank size, the prevalence of loans as earning assets and the relative amount of dividends. The adjusted R-squared is roughly 60 percent, showing that the chosen regressors do a good job at capturing variation in markups across banks, but within periods. Finally, when I add both year and bank fixed effects, we observe that all correlations are positive. The R-squared is now roughly 83 percent, meaning that the explanatory variables explain the vast majority of variation of markups across banks and across periods.

Columns 5 to 8 of Table 1.3 shows the correlations with the log-markdown on deposit rates. These correlate positively with bank size, the prevalence of loans as earnings assets and relative dividends within banks and within periods. Conversely, markdowns correlate negatively with leverage and profitability. The adjusted R-squared is just above 13 percent, suggesting again that the chosen regressors do a poor job at explaining variation in markdowns within banks and within years. Adding year fixed effects flips all signs, except for the ones on bank size and relative dividends. However, the adjusted R-squared jumps to roughly 76 percent. This means that the right-hand side variables explain most variation in markdowns within banks, but across years. Using instead bank fixed effects shows that markdowns are positively correlated with bank size, the prevalence of loans as earning assets and the relative dividends. They are negatively correlated with leverage and profitability. The adjusted R-squared is roughly 51 percent, indicating that the regressors explain only half of the variation of markdowns across banks, but within periods. Finally, when using both year and bank fixed effects, markdowns correlate positively with profitability and relative dividends. They instead correlate negatively with bank size, leverage and the prevalence of loans as earning assets. The adjusted R-squared is roughly 92 percent, suggesting that the chosen regressors capture almost all variation of markdowns across banks and years.

Overall, considering the last columns of each table, we can observe that bigger banks tend to charge higher markups but lower markdowns. This shows that smaller banks tend to exert their market power more on deposit markets rather than loan markets. We can draw similar conclusions regarding leverage: more leveraged banks charge higher markups and lower markdowns. Profitability correlates positively with both markups and markdowns. However, more profitable banks have higher markdowns than markups relative to less profitable banks. This suggests that markdowns on deposit rates play a larger role on the profitability of banks relative to markups on lending rates. Banks with a higher prevalence of loans as earning assets relative to held securities charge higher markups, but lower markdowns. In other words, banks relying relatively more on securities than loans as earning assets tend to charge higher markdowns on deposit rates. Finally, banks that pay more dividends relative to their interest income charge higher markups and markdowns.

1.4.3 Relationship with other measures of competition

How do markups and markdowns compare with existing measures of competition? To address this question, I compute the Herfindahl-Hirschman Index (HHI). The HHI is a measure of concentration based on market shares. It requires bank-level observations and yields an aggregate number within a set of banks. Analytically, it is computed as

$$\mathrm{HHI}_{gt} = \sum_{i \in g} s_{it}^2, \tag{1.10}$$

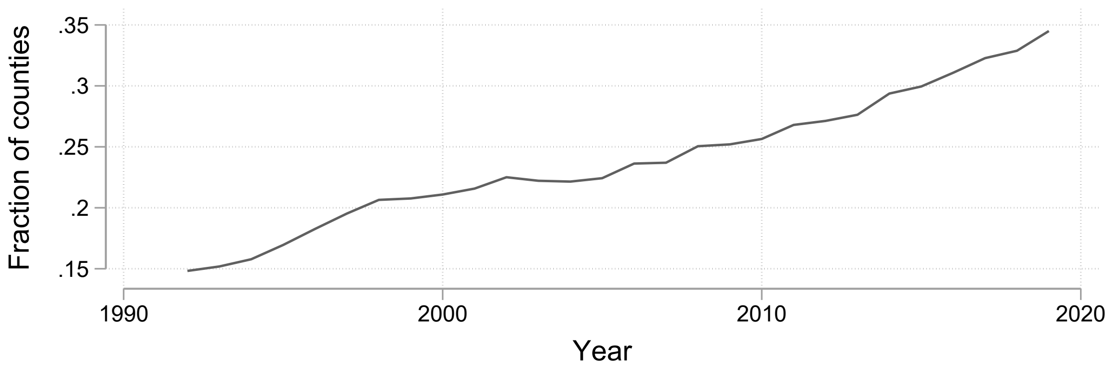

Figure 1.4: Fraction of counties that host only one bank, such that the HHI (on either loans or deposits) equals one.

for every period t, where g is a set of banks and s_{it} is the market share bank i has in period t within group g. In this formulation, the HHI ranges between $1/N_g$ and 1, where N_g is the number of banks in group g. When the HHI equals $1/N_g$, banks in group g have uniform market shares. This is usually associated with a highly competitive environment. Conversely, when the HHI equals one, there exists only one bank in group g that serves the entire market. This is usually thought of a highly monopolistic environment. As I have computed markups on lending rates and markdowns on deposit rates, I compute the shares s_{it} relative to both total net loans and total domestic deposits. I consider both US counties and states as delimiters that determine the sets g. However, there is a considerable number of counties in the US where only one bank operates. Figure 1.4 provides graphical evidence of this phenomenon. For all those counties, the HHI assumes it maximal value, one. For this reason, I present state-level evidence.

Figures 1.5 and 1.6 show the time series behavior of the average, median and interquartile range of the HHI computed on, respectively, loans and loans at the state level. These can be compared with Figures 1.2 and 1.3 respectively. The HHIs both have a slight upward trend over time. The cross-sectional variation in both indices is also slightly increasing over the years. The HHIs do not feature as much time series variation as the markups or markdowns. Markups tend to have a downward trend, while the HHI on loans trends upward. The HHIs being increasing over the sample period may be due to continued Mergers and Acquisitions (M&A) activity. In particular, M&A in the US banking industry are due to two main reasons. One is as part of recovery plans, often under the supervision or direction of the FDIC. The other is as part of deliberate deals for strategic reasons. The former reason increases the HHIs for mechanisms that are not related to market power but, rather, are due to financial stability concerns and to the intervention of the policy maker.

The graphical inspection is confirmed with Tables 1.4 and 1.5. The former shows the results from regressing the state median markup on the HHI on loans. The latter shows the results from regressing the state median markdown on the HHI on deposits. Particularly, regardless of whether I control for state or year fixed effects, markups correlate negatively with the HHI on loans. Conversely, markdowns correlate positively with the HHI on deposits. These results seem to suggest that the HHI on loans does not capture the same phenomenon as the markup on lending rates. The tables remain qualitatively unchanged if I compute the arithmetic average of markups or markdowns rather than the median.

The Boone (2008) indicator relies on measures of profitability and costs. Following Boone

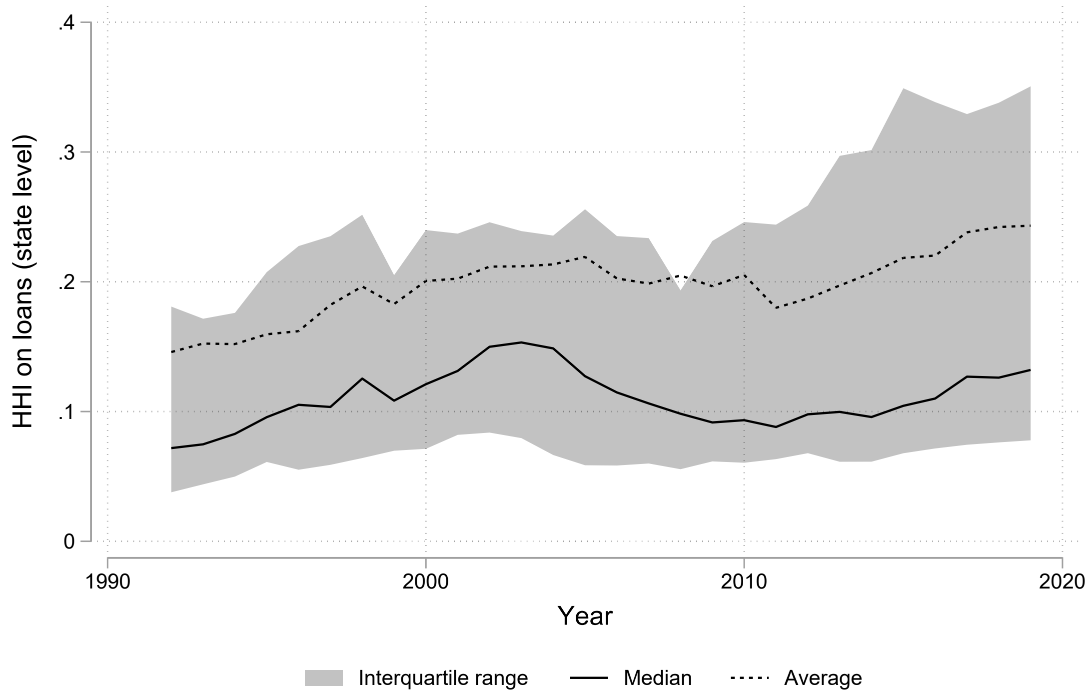

Figure 1.5: Average, median and interquartile range of the state-level HHI index on loans.

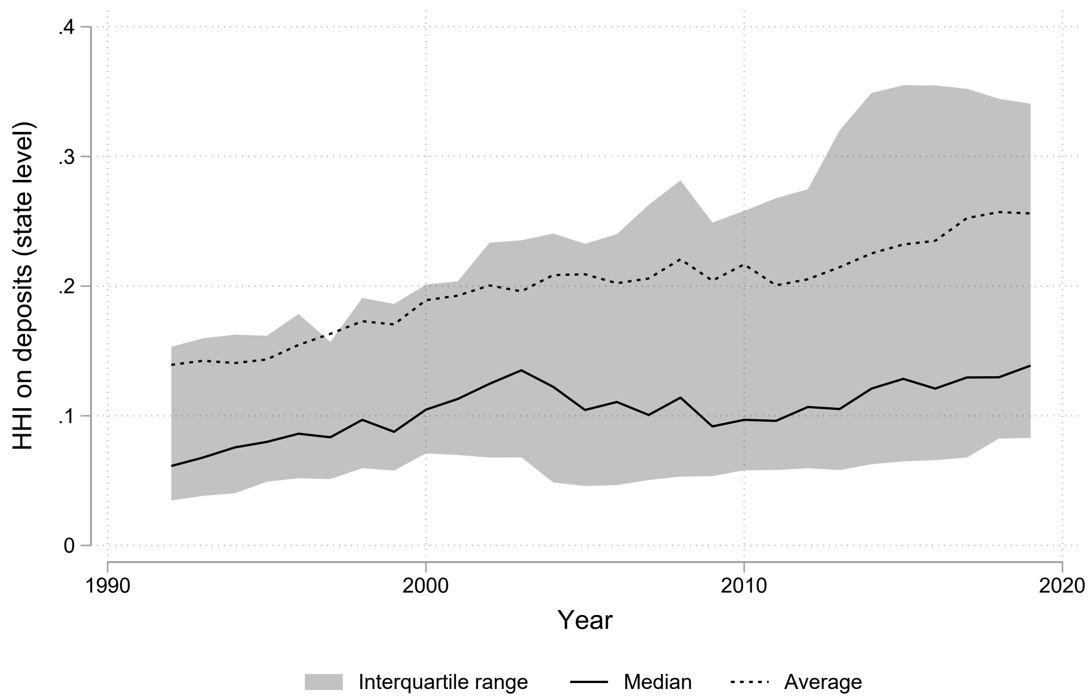

Figure 1.6: Average, median and interquartile range of the state-level HHI index on deposits.

Table 1.4: Results from regressing state-median markups on the HHI on loans. All variables are in logs.

	Log(Markups) — state-level medians			
	(1)	(2)	(3)	(4)
Log(HHI loans)	-0.0571^{***}	-0.1122^{***}	-0.0366^{***}	-0.0093
	(0.007)	(0.014)	(0.004)	(0.005)
Constant	0.4240^{***}	0.3072^{***}	0.4673^{***}	0.5252^{***}
	(0.017)	(0.030)	(0.008)	(0.010)
Observations	1559	1559	1559	1559
State FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Adjusted R^2	0.0370	0.1549	0.7784	0.9125

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.5: Results from regressing state-median markdowns on the HHI on deposits. All variables are in logs.

	Log(Markdowns) — state-level medians			
	(1)	(2)	(3)	(4)
Log(HHI deposits)	0.2376^{***}	0.5349^{***}	0.1252^{***}	0.0290^{**}
	(0.021)	(0.040)	(0.007)	(0.009)
Constant	0.6465^{***}	1.2934^{***}	0.4019^{***}	0.1923^{***}
	(0.051)	(0.090)	(0.018)	(0.020)
Observations	1559	1559	1559	1559
State FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Adjusted R^2	0.0741	0.1612	0.8898	0.9658

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

et al. (2005), I compute the Boone indicator using the following regression for every year t:

$$\log\left(\frac{\pi_{it}}{\pi_t}\right) = \kappa_{0,t} - \kappa_{1,t}\log\left(\frac{\mathrm{mc}_{it}}{\mathrm{mc}_t}\right) + \varsigma_{it}, \tag{1.11}$$

where π_{it} is the profit of bank i in period t and mc_{it} is the marginal cost. The terms π_t and mc_t are reference points within the cross-section. Ideally, they would correspond to the maximum. However, due to the presence of outliers and similarly to Boone et al. (2005), I use the 98$^{\text{th}}$ percentile. I estimate the marginal costs by fitting a translog cost function on the data. Specifically, I fit the following equation

$$\begin{aligned}\log(C_{it}) = \kappa_0 + \kappa_i + \kappa_t + \kappa_Q \log(Q_{it}) + \kappa_Q Q\log(Q_{it})^2 + \kappa_I \log(\mathrm{INTEXP}_{it}) + \\ + \kappa_W \log(W_{It}) + \kappa_O \log(\mathrm{OOEXP}_{it}) + \sum_{x_{it}\in\Omega_{it}}\sum_{y_{it}\in\Omega_{it}} \kappa_{xy}\log(x_{it})\log(y_{it}),\end{aligned} \tag{1.12}$$

where C_{it} is the sum of interest expenses, commission and fee expenses, trading expenses, personnel expenses, other administrative expenses, and other operating expenses, measured in millions of US dollars, Q_{it} is the quantity of output and is measured as total assets in millions of US dollars, INTEXP_{it} is total interest expense relative to total assets, W_{it} is total wages and salaries divided by total assets and OOEXP_{it} is other operating expenses over total assets and $\Omega = \{\mathrm{INTEXP}_{it}, W_{it}, \mathrm{OOEXP}_{it}\}$. From this, I can parametrically compute the marginal cost as

$$\mathrm{mc}_{it} = \frac{C_{it}}{Q_{it}}\left[\kappa_Q + 2\kappa_{QQ}\log Q_{it} + \kappa_{QI}\log(\mathrm{INTEXP}_{it}) + \kappa_{QW}\log(W_{it}) + \kappa_{QO}\log(\mathrm{OOEXP}_{it})\right]$$

Results about the Boone indicator will appear in future revisions of this work.

1.5 Measures of financial stability

Following Hillegeist et al. (2004), I consider three alternative measures of financial stability. Two are accounting-based and one is a market-based probability of bankruptcy. The accounting-based measures have been proposed by Altman (1968) and Ohlson (1980) respectively. The former is known as Z-score and the latter as O-score. They are the predicted values of reduced-form logit models of a future bankruptcy indicator on a set of balance sheet ratios. These measures rely only on the availability of balance sheet data. I can therefore compute them given the sample at hand. The market-based measure instead has been proposed by Merton (1974) and is commonly known as *distance to default*. In his approach, Merton considers a firm's equity as a call option on the company's assets using a structural asset pricing model. However, using this approach requires market data. Therefore, its applicability is restricted to banks that are publicly listed. Hillegeist et al. (2004) show that all three measures of financial stability are good predictors of a firm's bankruptcy. However, they also show that Merton's distance-to-default outperforms the Z-score and the O-score in terms of out-of-sample predictive power.

The Z-score and the O-score consist of the predicted values of two logit models. The dependent variable is an indicator, which equals one if a firm goes bankrupt in the two years ahead and zero otherwise. The independent variables are a set of balance sheet ratios. The two scores differ in the set of regressors. While Altman (1968) and Ohlson (1980) use logit models, I use linear probability models. This is driven by the scarcity of yearly bankruptcies in the data I use relative to the number of banks in each cross-section. Because of this, the maximum likelihood estimator for the logit parameters converges for neither the Z-score nor the O-score. Omitting the bank-year subscripts for notational convenience, the regression model I fit in order to compute the Z-score is

$$B = \delta_0 + \delta_1\frac{WC}{TA} + \delta_2\frac{RE}{TA} + \delta_3\frac{EBT}{TA} + \delta_4\frac{V_E}{TL} + \delta_5\frac{S}{TA} + u, \tag{1.13}$$

where B is the bankruptcy indicator, TA is total assets, TL is total liabilities, WC is working capital, RE is retained earnings, EBT is earnings before taxes, V_E is market value of equity and S is sales. I approximate the market value of equity with Tier-1 capital. The regression model I use to obtain the O-score instead is

$$B = \zeta_0 + \zeta_1 \frac{TL}{TA} + \zeta_2 \frac{WC}{TA} + \zeta_3 \frac{CL}{CA} + \zeta_4 \frac{NI}{TA} + \zeta_5 \frac{FFO}{TL} + \\ + \zeta_6 INTWO + \zeta_7 OENEG + \zeta_8 CHIN + v, \tag{1.14}$$

where CL is current liabilities, CA is current assets, NI is net income, FFO is pre-tax income plus depreciation and amortization, $INTWO$ is an indicator on whether cumulative net income over the previous two years is negative, $OENEG$ is an indicator on whether owners' equity is negative and $CHIN \equiv (NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$ is the scaled change in net income. The Z-score and the O-score are the in-sample predicted values from the regression models in Equation (1.13) and (1.14) respectively.

Figures 1.7 and 1.8 show the time series behavior of the average, median and interquartile ranges of the Z-scores and the O-scores. Z-scores have historically trended upwards, while O-scores did not. The cross-sectional dispersion of Z-scores has remained substantially constant, while that of O-scores has spiked with great recession. It is also useful to compare these figures to the time series of bankruptcies in Figure 1.1. We observe that the O-scores pick up the spike of bankruptcies around the great recession better than the Z-scores.

The distance to default measure is the difference between the asset value of the firm and the face value of its debt, scaled by the standard deviation of the firm's asset value. In his model, Merton (1974) models the firm's market equity value is implied by the following model:

$$\begin{aligned} V_E &= V_A N(d_1) e^{-dT} - X N(d_2) e^{-rT} + (1 - e^{-dT}) V_A \\ d_1 &= \frac{\log(V_A/X) + \left(r - d - s_A^2/2\right) T}{s_A \sqrt{T}} \\ d_2 &= d_1 - s_A \sqrt{T} \\ s_E &= \frac{V_A e^{-rT} N(d_1) s_A}{V_E}, \end{aligned}$$

where V_E is the market value of the bank, V_A is the value of the bank's assets, T is a maturity date, d is the dividend rate expressed in terms of V_A, r is the risk-free interest rate, X is the face value of debt maturing at time T, s_A is the volatility of the value of assets and s_E is the volatility of the value of equity. The system of equations above is solved numerically for V_A and s_A. Following Anginer et al. (2014), I map the remaining variables to the following data. The volatility of equity s_E is the standard deviation of daily equity returns over the past year. The risk-free rate r is the 1-year yield on US Treasury bills. The value of equity V_E is mapped to the market value and X is taken as total liabilities from balance sheets. I set the maturity T equal to one year. After having solved for V_A and s_A, the distance to default dd is computed as

$$dd = \frac{\log(V_A/X) + \left(m - d - s_A^2/2\right) T}{s_A \sqrt{T}},$$

where m is the asset return and is set to equal the equity premium. The results using distance-to-default will appear in future revisions of this paper.

1.6 Financial stability and market power

Similarly to Anginer et al. (2014), I regress measures financial stability on market power. The key difference here is the measure of market power. Rather than focusing on price-cost margins that

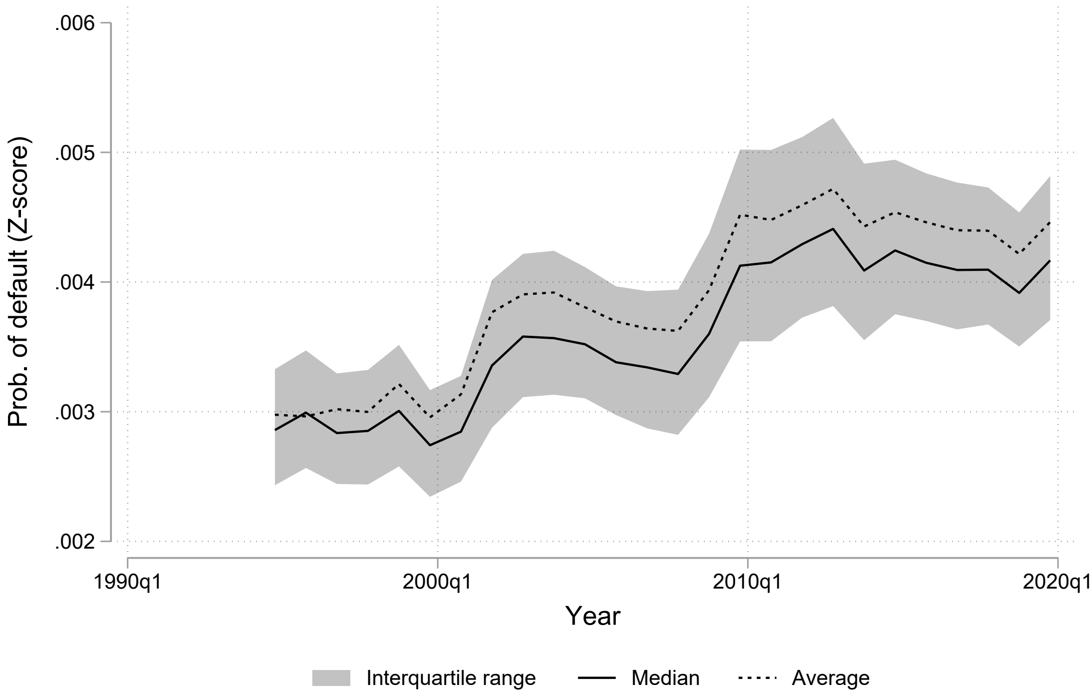

Figure 1.7: The average, median and interquartile range of the Z-scores.

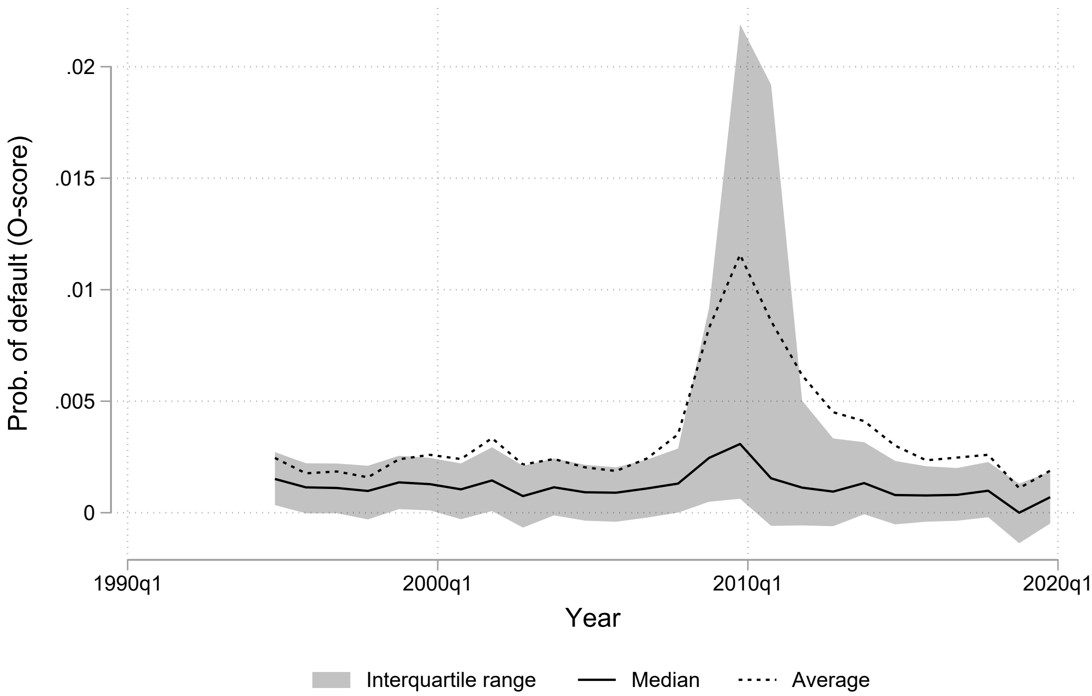

Figure 1.8: The average, median and interquartile range of the O-scores. Some probabilities are negative because they are the predicted values of a linear probability model.

relate to market power on loan markets, I disentangle a bank's market power as coming from two different markets: loan and deposit markets. In particular, I use the markup on lending rates as a measure of market power on loan markets. Conversely, I use the markdown on deposit rates as a measure of market power on deposit markets. The baseline regression model is

$$\text{Prob}\left(\text{bankruptcy}_{it}\right) = \gamma_0 + \gamma_i + \gamma_\mu \log(\mu_{it}) + \gamma_\psi \log(\psi_{it}) + \eta_{it}, \tag{1.15}$$

where μ_{it} is the markup on lending rates and ψ_{it} is the markdown on deposit rates. The left-hand side variable is either the Z-score or the O-score.

This work focuses on the *correlation* between market power and financial stability, rather than on causation. Both the right-hand side and the left-hand side variables in Equation (1.15) are computed from balance sheet data. One potential concern is that there will be spurious correlation. Because all data comes from balance sheet and income statement data, the regression may pick up mechanical correlation due to within-bank variation. However, all the variation in the measures of market power comes from income statement variables. Conversely, the left-hand side variables have been primarily obtained from balance sheet (not income statement) ratios. While income statement and balance sheet variables in levels are obviously correlated, adding bank fixed effects will capture their common variation within banks and leave variation across banks.

The main results are shown in Tables 1.6 and 1.7. The former uses the Z-score as dependent variable, while the latter uses the O-score. In all cases, markups on lending rates always correlate negatively with the scores. Markdowns on deposit rates correlate positively with both scores when considered alone, and negatively when holding markups constant. The inclusion of time fixed effects does not considerably change the point estimates. All correlations are highly significant. The magnitudes of the correlations differ, depending on whether one considers the Z-score or the O-score. A one percent increase in markups is associated with a decrease in the Z-score of 0.34 to 0.47 percent and a decrease in the O-score of 0.56 to 1.92 percent, depending on whether markdowns are held constant. Conversely, a one percent increase in markdowns is associated with an increase in the Z-score of roughly 0.11 percent and an increase in the O-score, or a decrease of 1–5 percentage points if markups are held constant.

1.7 Conclusion

In this paper I use the production approach to the estimation of markups and markdowns to banking data. I compute markups on lending rates and markdowns on deposit rates. I find that markups have generally been trending downwards over the years, while markdowns have been increasing, especially after the great recession. I correlate these new measures to observable bank characteristics. I find that bigger banks tend to charge a higher markup on lending rates and a lower markdown on deposit rates. Both markups and markdowns correlate positively with bank profitability. These two findings together suggest that smaller, more local banks have driven their profitability through deposit rates, particularly after the great recession. I compare the measures with the Herfindahl-Hirschman Index (HHI), which is a widely used measure of market concentration. I find that the HHI on deposits correlates with markdowns on deposit rates, although imperfectly. On the other hand, markups on lending rates correlate negatively with the HHI on loans. Finally, I relate the measures of markups and markdowns with measures of financial stability. Particularly, I estimate the Z-score and the O-score, which can be interpreted as default probability for each bank. I find that higher markups are associated with a lower probability of bankruptcy for banks. Conversely, markdowns are positively correlated with bankruptcy probability, but only if markups are not controlled for.

Table 1.6: Estimates of the coefficients for the model in Equation (1.15). The dependent variable is the log of the Z-score. All specifications include bank fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Log(markup)	-0.4042^{***}	-0.3465^{***}			-0.4111^{***}	-0.4730^{***}
	(0.002)	(0.002)			(0.003)	(0.003)
Log(markdown)			0.1170^{***}	0.1047^{***}	-0.0020	-0.1103^{***}
			(0.001)	(0.002)	(0.001)	(0.002)
Constant	-5.3579^{***}	-5.3877^{***}	-5.5897^{***}	-5.5874^{***}	-5.3558^{***}	-5.3042^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	148555	148555	148028	148028	147237	147237
Year FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.6294	0.6603	0.5898	0.6232	0.6316	0.6705

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7: Estimates of the coefficients for the model in Equation (1.15). The dependent variable is the log of the O-score. All specifications include bank fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Log(markup)	-0.5665^{***}	-1.0130^{***}			-1.9208^{***}	-1.8635^{***}
	(0.012)	(0.017)			(0.023)	(0.024)
Log(markdown)			0.0287^{***}	0.1026^{***}	-0.5461^{***}	-0.7145^{***}
			(0.004)	(0.011)	(0.008)	(0.015)
Constant	-5.8748^{***}	-5.6143^{***}	-6.2049^{***}	-6.2003^{***}	-5.1237^{***}	-5.1679^{***}
	(0.008)	(0.010)	(0.003)	(0.003)	(0.013)	(0.014)
Observations	137675	137675	137314	137314	136807	136807
Year FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.3141	0.3718	0.3016	0.3540	0.3384	0.3827

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This paper contributes to two main branches in the banking literature. First, to the best of my knowledge, this is the first paper that estimates markups and markdowns with banking data. While the methodology is not new, I adapt it to better suit the intermediation approach to production in banking. Under the assumption that markups and markdowns are measures of market power in loan and deposit markets respectively, this paper provides a new measure of competition for the banking industry. The main novelty here is that I disentangle market power on output markets (loans) from market power on input markets (deposits). The coexistence of market power on both inputs is not new to existing literature in Macroeconomics and Industrial Organization. However, this paper brings this concept to the banking industry. Second, I revisit the correlation between market power and financial stability. While most existing results find a positive relationship between market power and the probability of default for banks, I find that higher market power in lending markets is associated with lower bankruptcy probability, although the magnitude of the correlation is small. Instead, there is a positive relationship between markdowns and default probability, which vanishes once I control for markups.

Future revisions of this work will improve the paper in several directions. First, I will compare markups and markdowns with the Boone indicator. Comparing markups and markdowns to the Boone indicator will provide an important assessment for the measures I compute. However, the Boone indicator does not allow to disentangle market power on output or input markets. Second, I will add a new measure of financial stability. Merton's *distance to default* provides a more robust alternative to the Z-score and the O-score. The distance to default is based on the idea that a bank's equity can be seen as a call option on a bank's assets.

The main direction for future research consists in understanding the determinants of market power in the banking industry. While I disentangle markups and markdowns, I do not investigate what is determining them.

Chapter 2

European Unemployment and Discounts during the Great Recession

This chapter is based on joint work with Diego Comin, Riccardo Franceschin and Antonella Trigari.

2.1 Introduction

Aggregate data on labor market outcomes reveal a significant amount of differences across European countries. Unemployment rates differ both in levels and in volatility. We seek to explain the differences across Germany, France, Spain and Italy in terms of unemployment dynamics, with particular focus around the Great Recession. We do so by introducing shocks to discount factors.

We start with a standard, representative agent Diamond-Mortensen-Pissarides (DMP) model of labor market with search and matching frictions. In addition to the more traditional productivity shock, we augment the model with a discount factor shock. The role of discount factors in labor market outcomes is a recent addition to the literature. Discounts are considered a possible explanation of observed unemployment fluctuations.[1] We also briefly study the effect of a possible separation shock.

We first provide evidence that returns on European financial assets are highly correlated with unemployment across all countries we examine, possibly more than labor productivity. We then assess the ability of discount factors and workers' productivity to generate variation in unemployment. We find that discount factors are a promising source of variation to explain fluctuations in European unemployment.

We proceed by analyzing the predictions of the model through impulse-response functions. We consider how these predictions vary after changes in the calibration, which reflect changes in Labor Market Institutions (LMI). Changes to the average job-finding rate, the average separation rate and the extent of wage rigidity account for many differences between the US and EU labor markets. However, they do not account for the differences across EU markets.

We estimate exogenous shocks to the aggregate discount factor and aggregate productivity directly from the data. We input the shocks into the model and obtain simulations. We then compare the simulations to the data to assess the performance of the representative agent model in explaining observed data. We find that discount factor shocks can explain a significant part of the variation in unemployment for all European countries, contrary to productivity shocks, even without wage rigidity.

[1] See Hall (2017) and Borovička and Borovičková (2018).

This paper contributes to several branches of the literature. First, we document that discount factor shocks are a promising explanation for the volatility in European unemployment, similarly to Hall (2017) and Borovička and Borovičková (2018). We additionally propose a way to estimate discount factor shocks from stock market data, in a similar spirit to Borovička and Borovičková (2018). Second, we find that the representative agents DMP model with country-specific productivity and discount factor shocks cannot fully account for the differences in labor market outcomes across EU countries. A similar paper to ours is that of Albertini and Poirier (2014). They use a Bayesian Kalman Filter to estimate a DMP model where productivity and discounts are treated as unobserved states. We, instead, estimate externally the time series of productivity and discounts and use them to simulate the DMP model. In line with the results of Albertini and Poirier (2014), we find that discount factors plays a large role in explaining fluctuations of unemployment around the Great Recession.

The rest of this document is organized as follows. Section 2.2 presents the overall methodology we use. Section 2.3 explains the representative-agent search and matching model and inspects the main mechanisms with impulse-response functions. Section 2.4 presents the results we obtain with the representative agent model. Section 2.5 concludes.

2.2 Methodology

In this section we present the methodology we use to assess the effectiveness of a model in explaining observed variation. The overall procedure consists of simulating the model by "feeding in" exogenous shocks we estimate externally. This gives us simulated time series for labor market variables. We compare the simulations against the data. Our methodology is somewhat similar to that of Albertini and Poirier (2014). They estimate a DMP model with a Bayesian Kalman Filter and show the inferred series for productivity and discounts. These are treated as unobserved states in their estimation. In our paper, we estimate externally discount and productivity shocks and use them to simulate the DMP model. Now we proceed to explain how we obtain the time series estimates of the exogenous shocks.

2.2.1 Inference of SDF shocks

We rely on stock market data to obtain a time series for the discount factor, β_t. As the steps we take are applied to each national series independently, we omit country-specific indices in the notation that follows. It is important to note that we abstract from any microfoundation of the SDF and we are silent about the causes that move discounts. Our goal here is to find an observable proxy for the SDF.

Consider the following asset-pricing equation:

$$\mathbf{E}_t\left(\beta_{t+1}R_{t+1}\right) = 1, \tag{2.1}$$

where t denotes a month, β_{t+1} is the SDF and R_{t+1} is the gross return of a given financial asset from t to $t+1$. Log-linearizing (2.1) we obtain the relationship $\mathbf{E}_t(\hat{\beta}_{t+1}) = -\mathbf{E}_t(\hat{R}_{t+1})$, where the hat denotes that the variable is expressed in log-deviations from the steady state. By log-linearizing around the deterministic steady state, we are dropping any moment higher than the first. In the implementation that follows, we assume $\hat{\beta}_{t+1} = -\hat{R}_{t+1}$, making stronger assumptions about the relationship between the unobservable SDF and the observable returns.

As stock market returns exhibit much high-frequency variation, we smooth them by com-

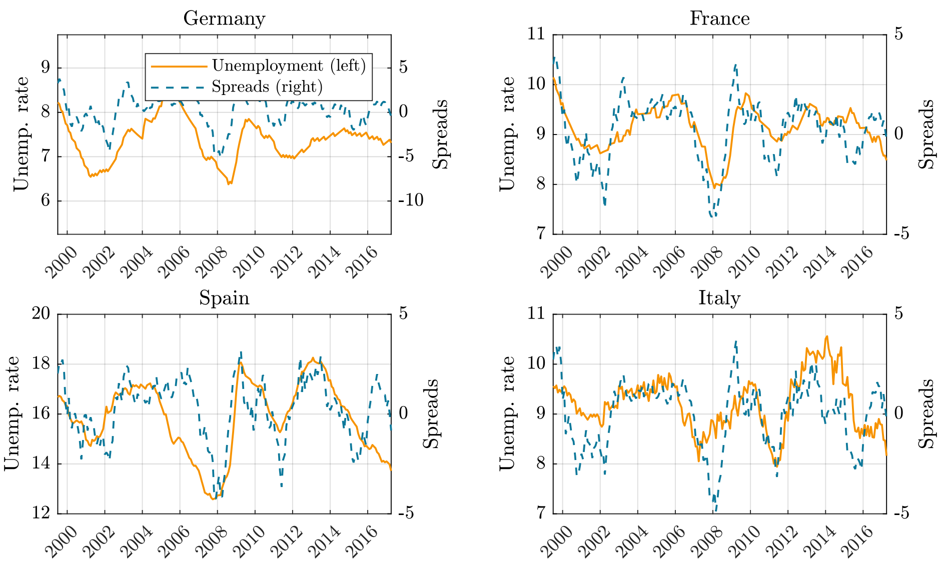

Figure 2.1: Unemployment (orange solid line) and the spread between stock market returns and the EONIA (blue dashed line), expressed as percent per month.

pounding returns in the following way:

$$1 + \bar{r}_t \equiv \sqrt[12]{\prod_{s=0}^{11} (1 + r_{t+s})},$$

where r_t is the monthly data point provided by WRDS. In words, we are taking the geometric average of a year of returns in a forward-looking way. Compounding returns forward reduces our sample size by one year at the end of the sample.

Because we solve the representative agent model by log-linearizing it, we do not relate levels of financial returns to the levels of the SDF. Instead, we relate their log-deviations from the steady-state. To this end, we construct the measure $\tilde{r}_t$ as

$$\tilde{r}_t = \log\left(1 + \bar{r}_t - r_t^f\right),$$

where we net out the risk free rate from the stock return of the financial asset, and we compute its trend-cycle decomposition using the Hodrick-Prescott filter with smoothing parameter $1600 \cdot 3^4$. Because we take logs, the resulting cycle can be interpreted as percent deviation from the trend. Figure 2.1 plots the measures $\tilde{r}_t$ together with observed de-trended unemployment for each of the four countries. The two series feature strongly correlated co-movements in each of the countries.

As we are constrained by data on productivity, which is available at quarterly frequency, we aggregate returns from monthly to quarterly. To compute the gross return for a given quarter, we compound the gross monthly returns observed within the quarter. The result scales to percent per quarter. Because of this transformation, we use the subscript t to indicate a quarter in the remainder of the paper.

In line with the asset pricing literature,[2] one may be worried that the risk premia we compute are not only driven by variations in discounts, but also in expected future cash-flows. In order to isolate variation in returns that we can attribute to discounts, we control for a measure of future

[2]Importantly, Campbell and Shiller (1988).

Table 2.1: Parameters for the quarterly process on β_t inferred from output per worker data. The steady state value $\tilde{\beta}$ is set and not estimated.

Parameter	Germany	France	Spain	Italy
$\tilde{\beta}$	0.9901	0.9883	0.9883	0.9955
ρ_β	0.74398	0.79455	0.7912	0.79725
σ_β	0.02733	0.02305	0.02371	0.02391

economic conditions. With US data, we could do so by controlling for dividend growth and/or variations in dividend-price ratios. However, as dividends in European markets do not play the same important role they do in US markets,[3] we use a different variable. The control variable we consider is the Leading Economic Indicator (LEI) by OECD, which provides qualitative forward-looking information about the state of the business cycle. This justifies the following specification for identification of SDF shocks:

$$\tilde{r}_t = \alpha + \rho_\beta \tilde{r}_{t-1} + \delta LEI_{t-1} + \eta_t. \tag{2.2}$$

By construction, the innovations η_t will not be systematically correlated with the Leading Economic Indicator. Hence we attribute the variation in these shocks to variation in discounts. We specify an AR(1) component in order to account for the dynamics we specify in the model. We use the estimates of the persistence ρ and the volatility of η_t to calibrate the parameters in Equation (2.17). We set the steady state value of the discount factor such that the associated discount rate equals the historical average of gross returns in the sample period. In order to simulate unemployment from the model, we feed $-\eta_t$ in place of $\sigma^\beta \varepsilon_t^\beta$ in Equation (2.17). The summary statistics of the regression are presented in Table 2.1.

In addition to the steps detailed above, we compute other measure of monthly SDF to assess the robustness of the methodology. The alternatives we consider are *(i)* using Euro Area-wide measure of LEI, as opposed to the country-specific one; *(ii)* inferring the process directly from stock market data, without accounting for the Leading Economic Indicators; and *(iii)* using only the part of variation of returns that could be predicted by dividend-price ratios or the LEIs. The results of this paper do not qualitatively change across the three alternatives. We also verify that European dividend-price ratios have low predictive power with respect to stock market returns.

2.2.2 Inference of aggregate productivity shocks

We employ a simpler, but similar, approach to obtain a series of productivity shocks to feed in the model. We use quarterly data on real GDP and on the number of employed people in each country to compute our measure of output per worker. We express the result as an index number, where the base period is the first quarter of 2010.

Similarly to before, we obtain log-deviations by computing the logarithm of productivity and then applying the HP filter with smoothing parameter equal to 1600. Figure 2.2 already showed the resulting series We finally fit an AR(1) process on the cycle component of the decomposition:

$$\tilde{z}_t = \omega + \rho_z \tilde{z}_{t-1} + \nu_t. \tag{2.3}$$

In order to simulate unemployment from the model, we feed ν_t in place of $\sigma^z \varepsilon_t^z$ in Equation (2.18). The summary statistics of the regression are presented in Table 2.2.

[3] We verify this with our data.

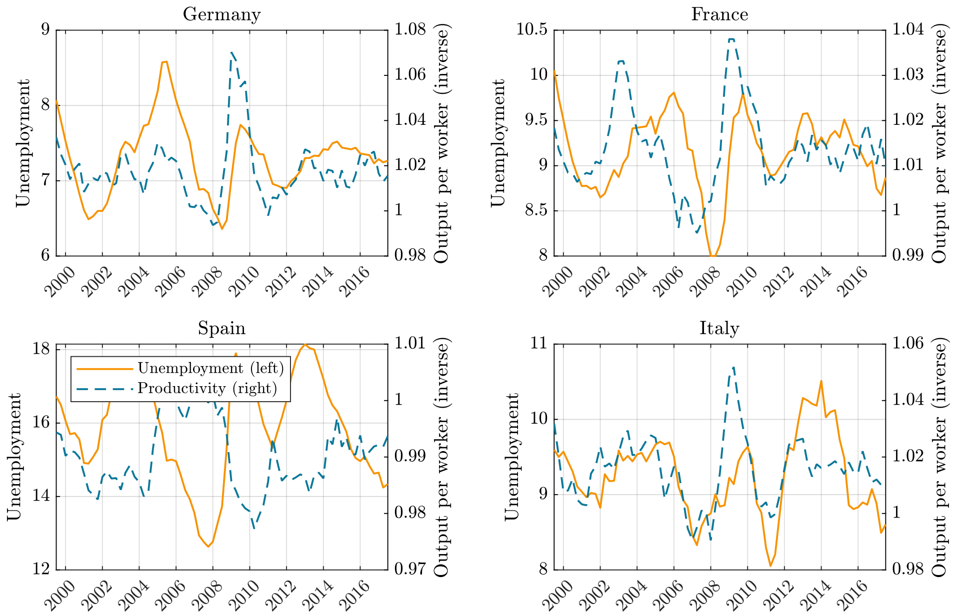

Figure 2.2: Detrended log of output per worker (blue dashed line, right axis) and detrended unemployment (orange solid line, left axis).

Table 2.2: Parameters for the quarterly process on z_t inferred from output per worker data. The steady state value $\tilde{z}$ is set and not estimated.

Parameter	Germany	France	Spain	Italy
$\tilde{z}$	1	1	1	1
ρ_z	0.82428	0.92073	0.96618	0.8597
σ_z	0.00850	0.00468	0.00371	0.0066

2.3 The Model

The model we use is a standard version of the Diamond, Mortensen and Pissarides (DMP) labor market model with search and matching frictions, whereby jobs are created according to the expected discounted profits over the match duration and exogenously destroyed at a given rate. We adjust our formulation to include three exogenous sources of variation: workers' productivity, an exogenous job destruction rate and a stochastic discount factor (SDF). In most of the analysis we focus on productivity and SDF shocks, but also briefly discuss separation shocks, as their impact is in part similar to SDF shocks.

While productivity and the separation rate are standard driving forces in the literature, the stochastic discounter only recently appeared in labor market models. We denote the SDF with β_{t+1}. We think of β_{t+1} simply as a random variable that allows agents to discount the future. In the consumption-based capital asset pricing model, the SDF is defined as the ratio of subsequent marginal utilities in consumption. In the financial economics literature, instead, the SDF is any random variable that prices a given asset. In line with Hall (2017), we abstract from any microfoundation, as we prefer to be agnostic about the microeconomic interpretation of a stochastic discounter. We let the SDF be time-varying to allow agents in our model to discount the future depending on the current aggregate state of the economy. We finally assume that the SDF is common across workers and firms.

Workers can be employed or unemployed and we abstract from labor force participation decisions. If unemployed, workers collect the unemployment benefit b and expect a future payoff stream by considering the probability p_t of finding a job. Such future payoff stream is discounted at the time-varying rate β_{t+1}. The sum of current and future payoffs gives the unemployment value, U_t:

$$U_t = b + \mathbf{E}_t \left\{ \beta_{t+1} \left(p_t W_{t+1} + (1 - p_t) U_{t+1} \right) \right\}. \tag{2.4}$$

If employed, workers earn the wage w_t and a future stream of wages that is discounted by β_{t+1} and consider the probability of job destruction s_t. The value of working is denoted with W_t and is given by:

$$W_t = w_t + \mathbf{E}_t \left\{ \beta_{t+1} \left((1 - s_t) W_{t+1} + s_t U_{t+1} \right) \right\}. \tag{2.5}$$

The difference between the value of working and the value of unemployment is the workers' surplus from employment:

$$W_t - U_t = w_t - b + \mathbf{E}_t \left\{ \beta_{t+1} (1 - s_t - p_t) (W_{t+1} - U_{t+1}) \right\}. \tag{2.6}$$

Firms hire workers by posting vacancies. If a firm hires, then it collects the value J_t, which is composed of the current profit, productivity minus wage, and the discounted future expected stream of profits:

$$J_t = z_t - w_t + \mathbf{E}_t \left\{ \beta_{t+1} \left((1 - s_t) J_{t+1} + s_t V_{t+1} \right) \right\}. \tag{2.7}$$

Posting a vacancy costs κ per period, but allows a firm to hire. The value of an open vacancy is given by:

$$V_t = -\kappa + \mathbf{E}_t \left\{ \beta_{t+1} \left(q_t J_{t+1} + (1 - q_t) V_{t+1} \right) \right\}, \tag{2.8}$$

where q_t is the vacancy-filling rate. Free entry drives the value of a vacancy to zero:

$$-\kappa + \mathbf{E}_t \left\{ \beta_{t+1} q_t J_{t+1} \right\} = 0 \tag{2.9}$$

$$\frac{\kappa}{q_t} = \mathbf{E}_t \left\{ \beta_{t+1} J_{t+1} \right\}. \tag{2.10}$$

By combining the value of a job J_t and the free-entry condition, we obtain:

$$J_t = z_t - w_t + \mathbf{E}_t \left\{ \beta_{t+1} \left(1 - s_t\right) J_{t+1} \right\}. \tag{2.11}$$

Workers and firms are matched according to a matching function m_t that we assume to be Cobb-Douglas:

$$m_t = \sigma^m u_t^\sigma v_t^{1-\sigma}, \tag{2.12}$$

where σ^m denotes the efficiency of the matching process, u_t is the unemployment rate and v_t is the vacancy rate. Unemployment at date $t+1$ equals date t unemployment plus exogenous layoffs, minus new matches:

$$u_{t+1} = u_t + s_t \left(1 - u_t\right) - m_t. \tag{2.13}$$

The probability for a worker to find a job must equal the number of new matches relative to the mass of unemployed workers, $p_t = m_t/u_t$; similarly, the probability for a firm to fill a vacancy is $q_t = m_t/v_t$.

The wage in this model is set according to the Nash bargaining protocol, whereby workers and firms agree on a wage that maximizes a function of the parties' surpluses:

$$w_t^{NB} = \arg\max_{w_t} \left(W_t - U_t\right)^\eta (J_t)^{1-\eta}. \tag{2.14}$$

The first-order condition for this problem gives the equilibrium wage, which is determined by a surplus sharing rule:

$$w_t^{NB} = \eta \left(z_t + p_t \frac{\kappa}{q_t} \right) + (1 - \eta)\, b. \tag{2.15}$$

When we consider wage rigidity, we impose a rule such that

$$w_t = (1 - \gamma) w_t^{NB} + \gamma \bar{w}, \tag{2.16}$$

where $\bar{w}$ is the steady state value of the wage and γ is a parameter governing the degree of wage rigidity.

We close the model by introducing the stochastic processes for the exogenous variables. We specify AR(1) processes for each of them, which is common practice in the literature in order to introduce persistency effects in agents' expectations.

$$\log(\beta_t) = (1 - \rho^\beta) \log(\tilde{\beta}) + \rho^\beta \log(\beta_{t-1}) + \sigma^\beta \varepsilon_t^\beta, \tag{2.17}$$

$$\log(z_t) = (1 - \rho^z) \log(\tilde{z}) + \rho^z \log(z_{t-1}) + \sigma^z \varepsilon_t^z, \tag{2.18}$$

$$\log(s_t) = (1 - \rho^s) \log(\tilde{s}) + \rho^z \log(s_{t-1}) + \sigma^s \varepsilon_t^s, \tag{2.19}$$

where each of the shocks ε_t^i, with $i \in \{\beta, z, s\}$, is independently and identically distributed according to standard Gaussian distributions.

2.3.1 Calibration

As anticipated above, we start our analysis with a baseline monthly calibration that targets US labor market moments. We pick the calibration in Shimer (2005) as our baseline. This calibration represents a widely known benchmark for the literature. Table 2.3 presents the calibration. We normalize the average labor productivity to one. The unemployment benefit b is set to 0.4: this means that the unemployment benefit is roughly 40 percent of the average labor income, which

Table 2.3: Values of calibrated parameters expressed in monthly terms.

Target/Parameter	Meaning	Values
$\tilde{z}$	Steady-state value of productivity	1 (normalization)
b	Unemployment benefit	0.4
η	Workers' bargaining power	0.5
$\tilde{p}$	Target job-finding rate	0.45
$\tilde{q}$	Target vacancy-filling rate	0.7
σ^m	Matching efficiency	1 (normalization)
σ	Elasticity of matching to unemployment	0.5
$\tilde{s}$	Average job destruction rate	0.03
ρ^β	Persistence of SDF process	$0.95^{1/3}$
ρ^z	Persistence of productivity process	$0.95^{1/3}$
ρ^s	Persistence of separation rate	$0.95^{1/3}$
σ^β	Volatility of shocks to SDF	0.1527
σ^z	Volatility of shocks to productivity	0.015
σ^s	Volatility of shocks to separation rate	0.2887

amounts to approximately 0.96 with this calibration. We set the average separation rate s to 0.03, so that employment lasts roughly 2.7 years on average (33 months). We let the vacancy cost κ vary to target an average job-finding rate of 0.45 in US data and normalize the matching efficiency σ^m to one. We set the elasticity of matches to unemployment σ to 0.5, a midpoint of the estimates in the literature.[4] We set the worker's bargaining power η to 0.5 assigning equal power to both parties and satisfying the Hosios (1990) efficiency condition. The persistencies of the exogenous processes ρ_β, ρ_z and ρ_s are set equal in order to compare the Impulse-Response Functions that follow. Finally, we set the volatilities for the exogenous shocks σ^β, σ^z and σ^s so that the implied volatility of output, with each of those shocks alone, matches the observed volatility in the data. This implies that the Impulse-Response Functions should be interpreted relative to output.

We then develop our own calibration in order to assess the role of Labor Market Institutions. We do so by using the baseline calibration and changing the unemployment benefit b, the job-finding probability $\tilde{p}$ and the separation rate $\tilde{s}$ on a country by country basis.

To set a value of b, we use annual data on Net Replacement Rates (NRRs) by OECD. These measure the fraction of the average income that a household retains after a transition from employment to unemployment. The available data is rich in terms of slicing the reference population. We consider the NRRs for households composed of two adults with two children and where the second adult is inactive. We further narrow the choice of the value to those households that are two months into unemployment. As OECD provides an annual time series for the NRRs, we compute the historical average on the sample period we consider and we set this value to b in the calibration. We do not choose NRRs for households where the second adult is employed because the NRR, by definition, is considerably driven up by his/her income earnings.[5] This is

[4] See Blanchard and Diamond (1989) and Petrongolo and Pissarides (2001).

[5] In fact, for any given year in the OECD' dataset,

$$\text{NRR} \equiv \frac{y_{OW}}{y_{IW}},$$

where y_{OW} is out-of-work net household earnings and y_{IW} is in-work net household earnings. The two measures are taken after and before (respectively) the transition to unemployment. As both measures are net *household* earnings, both include any labor income earning that is got by the adult that does not transition to unemployment.

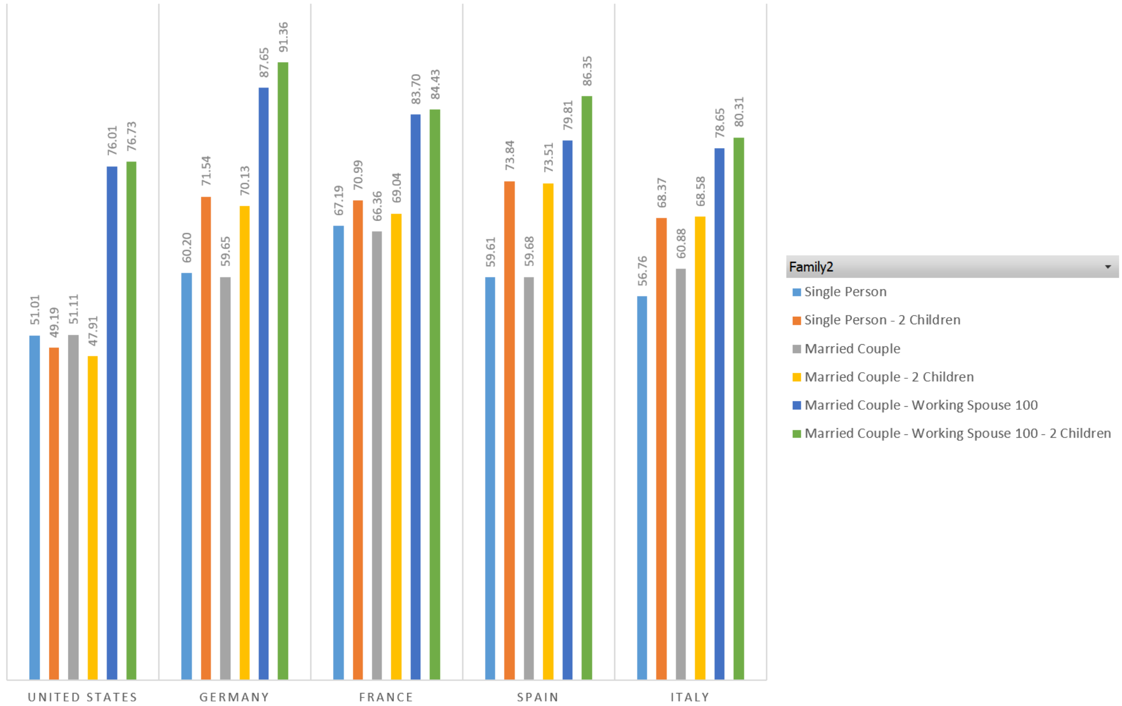

Figure 2.3: Net Replacement Rates by household composition. The values are averages of the yearly observations.

documented by Figure 2.3, where we also observe that the US generally provide lower benefit and assistance to unemployed households relative to European countries. Figure 2.4 shows the rates for the household composition we choose, by unemployment duration. We note that the levels of the NRRs tend to drop considerably in the long term (5 years). Given that the average duration of unemployment in European countries is roughly between 11 and 19 months,[6] and thus closer to two months than five years, we restrict our attention to the NRR measured at the second month of unemployment. We also see that the speed of the drop varies significantly across countries. Furthermore, we choose higher values of the NRR because OECD only includes cash flows in the calculation of the NRRs, omitting non-cash benefits. In the model, b represents any benefit a household might collect every period, including any non-monetary flow (e.g., home production, leisure). We therefore prefer picking the higher values of NRR.

We estimate the values of the steady state job-finding probability $\tilde{p}$ and the separation rate $\tilde{s}$ by partially replicating Elsby et al. (2013). The replication is necessary to extend their methodology to our sample period. Their results stop at 2009, while our sample period ends in August 2017. Following their steps, we compute the job-finding probabilities conditional on the duration of unemployment (less than a month, less than three months, less than six months and less than a year). Elsby et al. proceed to compute a set of optimal weights to average out the conditioning of each measure. In our replication exercise, we observe that their results are almost entirely driven by the job-finding probability for those who have been unemployed by less than a year. We therefore pick this duration of unemployment as representative of the unconditional job-finding probability. With such probability and with data on unemployment, Elsby et al. invert the continuous time-based law of motion of unemployment to recover the separation probability. We do the same here. Figure 2.5 shows the results we obtain by replicating Elsby et al. (2013)

[6] See Table 2.4 below. In particular, the average duration of unemployment is given by $1/\tilde{p}$. As we calibrate by targeting monthly moments, the average duration is expressed in months.

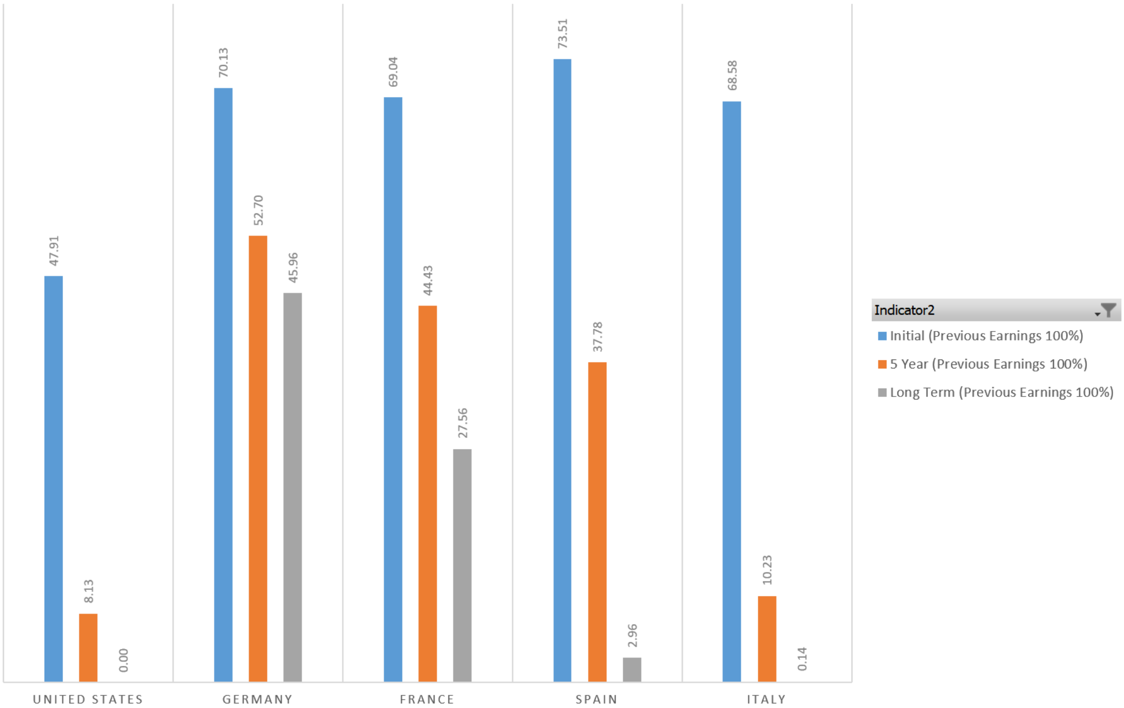

Figure 2.4: Net Replacement Rates by unemployment duration for married couples with two children and inactive spouse. The values are averages of the yearly observations. The data labeled with “5 year” are averages of the NRRs reported across durations. The data labeled with “long term” refer to households who have been unemployed for five years.

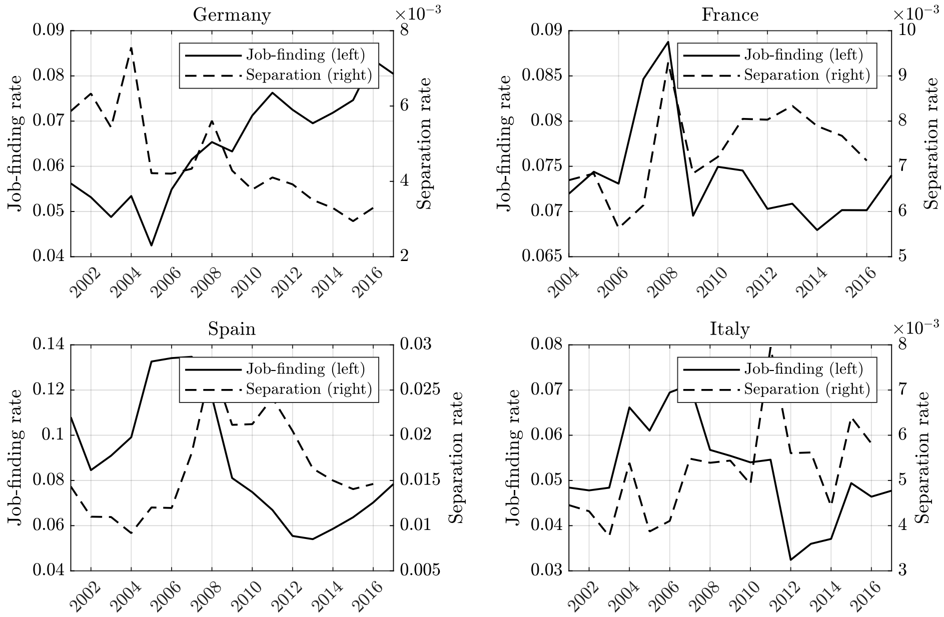

Figure 2.5: Job-finding and separation probabilities using the methodology in Elsby et al. (2013) on our sample period.

Table 2.4: Country-specific calibration.

Target	United States	Germany	France	Spain	Italy
b	0.4791	0.7013	0.6904	0.7351	0.6858
$\tilde{p}$	0.3559	0.0647	0.0740	0.0885	0.0519
$\tilde{s}$	0.0338	0.0045	0.0074	0.0164	0.0052
$\tilde{u}$	0.0603	0.0657	0.0906	0.1563	0.0908

on our sample period. We verify that our results largely coincide with theirs where the sample periods intersect. As their methodology gives annual estimates of the two probabilities, we take historical averages to set the steady state values $\tilde{p}$ and $\tilde{s}$.

With given values of the steady state transition probabilities, our model pins down the steady state values of unemployment through the steady state version of the law of motion of unemployment:

$$\tilde{u} = \frac{\tilde{s}}{\tilde{s} + \tilde{p}}. \tag{2.20}$$

Table 2.4 summarizes the values we set in our calibration. As we apply this calibration methodology also to US data, we can compare US steady state values with the corresponding European ones. Both the job-finding and the separation rates are significantly lower in the European countries we consider relative to the US. This implies both a longer average duration of unemployment (through lower $\tilde{p}$) and a longer average duration of employment (through lower $\tilde{s}$). Because of these differences, we refer to the US as a *fluid* labor market and to the European ones as *sclerotic*. In other words, fluid environments feature more faster transitions into and from unemployment relative to sclerotic ones.

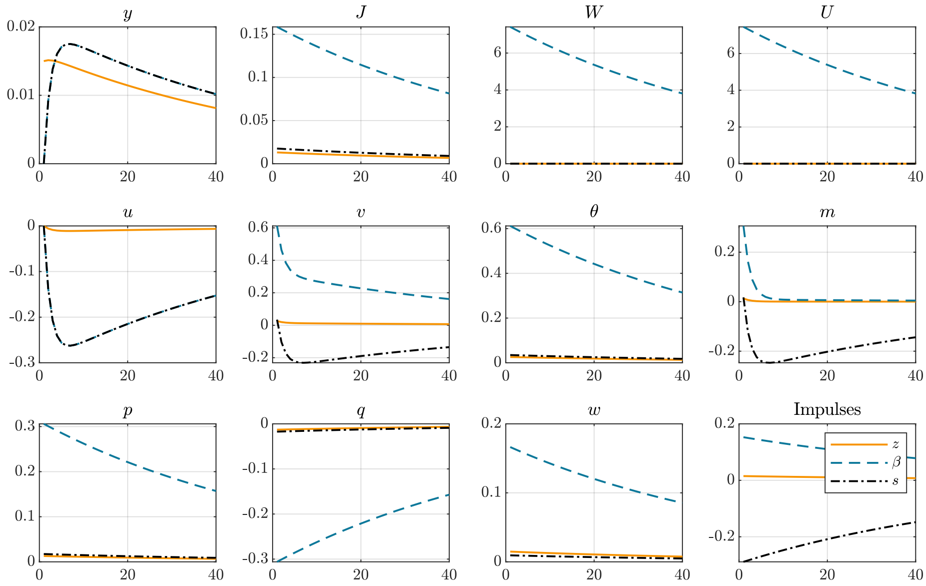

Figure 2.6: Impulse-Response Functions (IRFs) under the baseline calibration.

We also observe that the unemployment benefits differ from the baseline calibration. On average, European countries provide higher transfers to unemployed households than the US. As is known in the literature, unemployment benefits may play an important role in explaining unemployment fluctuations. For example, Hagedorn and Manovskii (2008) show that with high enough benefits and for particular values of the workers' bargaining power, a DMP model may not need wage rigidity to explain unemployment only through variation in workers' productivity.

Finally, we change the degree of wage rigidity. As mentioned above, we do so by setting values of γ in Equation (2.16). Setting $\gamma = 1$ means allowing for full flexibility in the wage bargaining protocol, while imposing $\gamma = 0$ pins down wages to their steady state value forever. While we do not calibrate the degree of wage rigidity, we change its value to arbitrary values to show how exogenous shocks differently propagate throughout the labor market.

As we anticipated above, we produce quarterly simulations. Therefore we also convert the monthly calibration to a quarterly one, specifically the average job finding and job separation rates.

2.3.2 Inspecting the mechanisms

We explore the qualitative predictions of our model using Impulse-Response Functions (IRFs). Figure 2.6 shows the Impulse-Response Functions of our model to shocks to the three exogenous variables of one standard deviation size. In particular, as mentioned above, the calibration of those standard deviations are such that a standard deviation of output simulated with each shock alone matches the data. The qualitative implications of the model are standard when compared to the literature. As already pointed out in Shimer (2005), productivity shocks cannot produce amplification of unemployment and number of vacancies relative to output. Consistently with the literature, shocks to the separation rate do not generate the negative correlation between unemployment and vacancies (also known as the Beveridge Curve).

Note that the impulse responses of output and unemployment are exactly the same in case of separation and discount factor shocks. This happens for two reasons. First, the processes are calibrated in such a way that the volatility of output is the same after each shock, separately, hits the economy. At the same time, the model assumes that output is unaffected by the two shocks upon impact and that it reacts only in subsequent periods. Second, both shocks enter discounting the same way—$(1 - s_{t+1})\beta_{t+1}$—hence the impact of these shocks on the value functions is similar. The difference is that only discount factor shocks enter the job creation condition while only separation shocks enter the law of motion of unemployment. This also explains the different responses in the evolution of vacancies.

2.3.3 The effects of SDF shocks vs productivity shocks

A positive shock to the discount factor enters the model through the firms' incentive to hire by making them more forward-looking. In other words, payoffs further ahead in the future are discounted less. This incentivizes firms to hire, raising vacancies and reducing unemployment. As more firms enter the market, total production increases, but only after one period (that is, not on impact). This happens because the model's timing implies that it takes one period for a new match to start producing. Unemployed workers find jobs more easily because of increased opportunities. At the same time, higher entry by firms makes it more difficult for each firm to find a worker. As the total surplus in the economy rises, wages rise. Compared to the shocks to productivity, shocks to the discount factor cause larger movements in labor market activity (vacancies, unemployment, job finding and job filling rates) relative to output. Moreover, movements in discounts can generate the Beveridge curve.

A positive shock to workers' productivity generates the same fluctuations in terms of sign of the discount shock. More firms enter the market and, as the overall surplus increases, wages rise. Job filling rates decrease for firms, while unemployed workers have better chances to find a job. The intuition for the effects is similar as the one for SDF shocks. A positive increase in workers' productivity also increases the firms' value of a job. However, this occurs because of higher current and future expected cash flows $z_{t+s} - w_{t+s}$ from the match, as opposed to higher valuation of future cash flows. Because of increased time t productivity, output responds on impact. More firms enter the market and, as the overall surplus increases, wages rise. Market tightness increases for firms, while unemployed workers have better chances to find a job. However, productivity shocks enter firms' value through the per-period surplus $z_t - w_t$, while discounts have multiplicative effects.

The effects of SDF and productivity shocks share some similarities. Both innovations enter the model by increasing the firms' incentive to hire. Discounts increase the value of a job by increasing its present value. Productivity, instead, increases the each period's cash flow. As time-t productivity immediately influences time-t output, the effect of z_t on output is non-zero on impact. Conversely, discounts have immediate effects on the vacancy-filling rate through the (log-linear) free-entry condition $\mathrm{E}_t(\hat{\beta}_{t+1} + \hat{J}_{t+1}) = -\hat{q}_t$, but do not immediately affect output or unemployment. Finally, while discount factors lower unemployment and therefore increase total production by entry of new firms, productivity shocks increase each firms' production.

The amplification of SDF shocks largely depends on the persistence of the SDF shocks and the extent of wage rigidity. The left panel of Figure 2.7 illustrates the point. For a given degree of wage rigidity, a decrease in the persistence of the SDF shock makes unemployment react in a much less volatile manner. Moreover, the role of wage rigidity in the amplification of the shocks changes depending on the persistence. We draw similar conclusions about productivity shocks, as illustrated on the right panel of Figure 2.7. It remains true, however, that productivity shocks generate variation of unemployment (relative to output) one order of magnitude lower than SDF shocks (as illustrated by the different scale of the two panels).

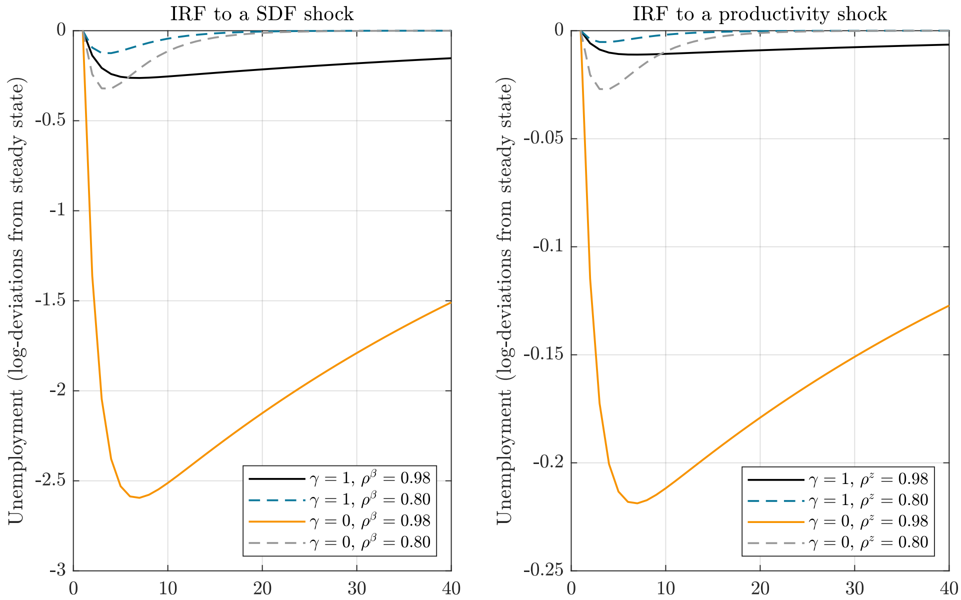

Figure 2.7: IRF of unemployment to a SDF shock (left) and to a productivity shock (right), for different wage rigidity (γ) and persistence of each shock (ρ^β, ρ^z).

2.3.4 The role of Labor Market Institutions

We begin analyzing the role of Labor Market Institutions by comparing the IRFs to the different shocks under different calibrations. As we are ultimately interested in the dynamics of unemployment, we focus on the response of unemployment to the different shocks and we provide the intuition for the changes by looking at the equations of the model.

We clarify here that we use the term “Labor Market Institution” in a broad sense. Through the lens of our model, a direct way a policy maker may influence labor markets is to change the policies to allocate unemployment insurance. However, we also think of LMIs as the environment in which the labor market exists. This includes, for example, the laws that define and regulate labor contracts. In this sense, LMIs also have an effect on how dynamic a market is, particularly in terms of the average durations of employment and unemployment.

The left panel of Figure 2.8 shows the response of unemployment to a positive discount factor shock calibrated with the AR(1) properties as in Table 2.3. However, the unemployment benefit, the job-finding probability and the separation rate are changed to capture a fluid labor market (the US) and a sclerotic labor market (European countries). In particular, the “fluid” calibration has $b = 0.4$, $\tilde{p} = 0.45$ and $\tilde{s} = 0.03$, which are the baseline values. The “fluid” (high b) calibration has $b = 0.7$ and $\tilde{p}$ and $\tilde{s}$ as above (where 0.7 approximates the values in European countries from Table 2.4). The “sclerotic” calibration has $b = 0.7$, $\tilde{p} = 0.07$ and $\tilde{s} = 0.008$ (again see Table 2.4).

We make two observations. First, unemployment benefits do not impact the transmission or amplification of SDF shocks, while they significantly amplify productivity shocks. The relative average value of non work to work activities—b in the model (with z normalized to 1)—has received a lot of attention in the literature.[7] This because the literature on unemployment dynamics within search and matching models has so far focused on productivity shocks as a driving force. Productivity shocks impact hiring by changing current and future cash flows, whose response is in turn largely determined by the relative value of b to productivity (via its effects on the relative

[7] See in particular Hagedorn and Manovskii (2008) and Chodorow-Reich and Karabarbounis (2016).

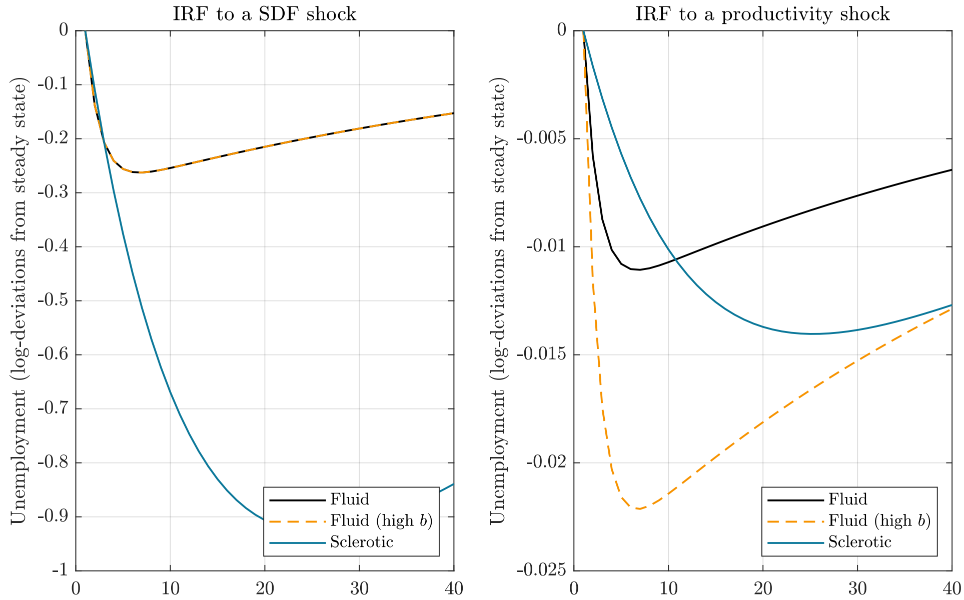

Figure 2.8: IRF of unemployment to a SDF shock (left) and to a productivity shock (right), for different calibrations.

values of productivity and wages). Discount shocks instead affect hiring by changing the valuation of given cash flows, in multiplicative manner, and their impact is thus unaffected by the relative average values of the cash flows components. Second, sclerotic labor markets exacerbate the effects of discount factor shocks on unemployment relative to fluid (with high b) markets: the response of unemployment is larger and more persistent. To understand why this is the case, consider the law of motion for unemployment (2.13) rearranged and log-linearized (assume the separation rate is constant):

$$\hat{u}_{t+1} = (1 - \tilde{s} - \tilde{p})\hat{u}_t - \tilde{p}\hat{p}_t.$$

Now, in fluid labor markets both $\tilde{p}$ and $\tilde{s}$ tend to be high, so that $1-\tilde{s}-\tilde{p}$ tends to be low. This means that the variation in unemployment is primarily driven by the job-finding rate. Conversely, in sclerotic labor markets, $\tilde{p}$ and $\tilde{s}$ are low, so that $1-\tilde{s}-\tilde{p}$ is high. This means that it is the variation in unemployment *growth* that is primarily driven by $\hat{p}_t$, which generate more persistent dynamics for unemployment.

The right panel of Figure 2.8 plots the response of unemployment to a positive productivity shock. Setting a high unemployment benefit in a fluid labor market amplifies the response of unemployment to a productivity shock, as discussed above. On the other hand, sclerotic markets increase the average duration of employment and unemployment, increasing the persistence of the response of unemployment to productivity shocks, but decreasing amplification.

We finally observe that the effects of Labor Market Institutions depend on how persistent the shocks are. The effect of the interaction between LMIs and the persistency of the shocks is different for SDF and for productivity impulses. Figure 2.9 documents this fact. Again, the two calibrations only differ because of different values of the transition probabilities $\tilde{p}$ and $\tilde{s}$. In the left panel we see that a persistent discount factor shock is greatly amplified by sclerotic environments relative to fluid ones, although the effect relies on the persistence of the shock. With less persistent shocks, discount factors are less amplified. In this case, the magnitude of the response of unemployment is roughly unchanged across calibrations, although its persistence is

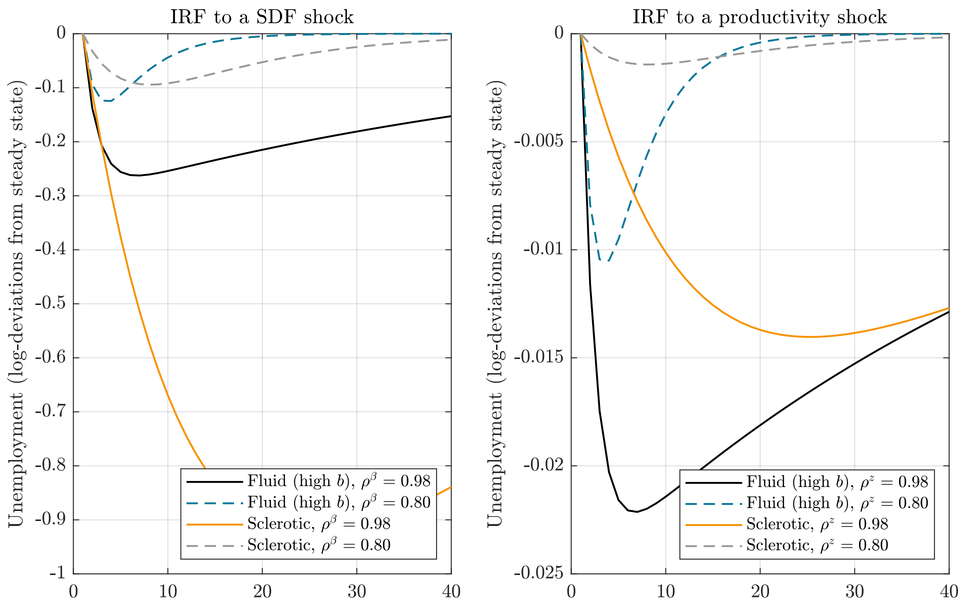

Figure 2.9: IRF of unemployment to a SDF shock (left) and to a productivity shock (right), for different calibrations and persistency of the shocks.

higher in sclerotic environments. The effect travels through the increased average duration of both employment and unemployment. Conversely, the persistency of productivity shocks is less crucial than the fluidity of the market for the amplification mechanism.

2.4 Results

In this section we illustrate the results we obtain with the methodology and the models presented above. While we do have results with the representative agent model, we are still working on the model with heterogeneous agents. At the current stage, we only present the results we obtain with the representative agent model, which justify our attention to dual labor markets.

We finally turn to generating the series of simulated unemployment. We obtain the simulations by feeding in the shocks as estimated in (2.2) (changing the sign) and in (2.3) into (2.17) and (2.18) respectively. We produce quarterly simulations.

First we show the simulations by only feeding in SDF shocks. We show how the simulations are affected by different degrees of wage rigidity and by sclerotic labor markets. We repeat the analysis with simulations obtained by only using productivity shocks. For these specific simulations, where we comment on the differences between fluid and sclerotic environments, we only vary the transition probabilities $\tilde{p}$ and $\tilde{s}$. This means we keep the relative value of non-work to work activity, b, pinned down to 0.4. We do so because we want to focus on the effect of slower transitions to and from unemployment and we want to abstract from different values of b. We also allow full wage flexibility. We finally allow both shocks into the final simulation, where we assess the relative contribution of each source of variation. As a way to numerically assess the "fit" of the simulations to the data, we regress simulated unemployment on observed (HP-filtered) unemployment. If the model is able to perfectly match the data, the slope of the regression will be unity. If the simulated variation is less than observed volatility, then the absolute value of the slope will be between zero and one. If the simulations are more volatile then the data, the

Table 2.5: Covariance between simulations (by wage rigidity) and data relative to the volatility of observed (HP-filtered) unemployment. Only SDF shocks

Wage rigidity	Germany	France	Spain	Italy
Flexible ($\gamma = 1$)	0.3415	0.7627	0.1523	0.3087
Semi-rigid ($\gamma = 0.5$)	0.4284	1.0123	0.2071	0.3893
Rigid ($\gamma = 0$)	0.5571	1.4315	0.3020	0.5133

absolute value of the slope coefficient will be greater than one. If the sign of the slope is negative, then positive variation in the data is associated with a negative variation in the simulations.

Figure 2.10 shows the simulations obtained by only using SDF shocks by degree of wage rigidity. In doing this, we completely shut down productivity shocks. We observe that wage rigidity amplifies the variation of unemployment, although the effect is different across countries. This is not surprising, as we verified with the IRF in Figure 2.7 that the effect of wage rigidity varies with the persistence of discounts. The persistence of discounts in our data is between 0.7 and 0.8. In particular, the persistence in Germany is lower than in other countries, explaining why the effect of wage rigidity in Germany is weaker. Table 2.5 accompanies these findings. We observe that introducing wage rigidity increases the correlation between the simulations and the data, with the effect being weaker in Germany. In the case of France, full wage rigidity makes the simulations more volatile relative to the data.

Our methodology presents timing issues, which are highlighted by the increased persistence of unemployment under the sclerotic calibrations. These are especially noticeable in Spain and in Italy. Overall we observe that our simulations lag relative to the data by roughly two to five quarters, depending on the country. We also cannot explain the peak in unemployment in Germany in 2005.

Figure 2.11 shows the simulated unemployment using only SDF shocks, by fluidity of the labor market. Here, the unemployment benefit b is set to 0.4 to focus on the differences caused by the variation in transition probabilities. As we observed in the Impulse-Response Functions in Figure 2.8, fluid labor markets allow for similarly volatile but less persistent responses of unemployment relative to sclerotic environments. Moreover, the (small) differences between fluid and sclerotic environments are consistent with the finding in the left panel of Figure 2.9, where we showed that SDF shocks with lower persistence are less amplified in sclerotic markets than highly persistent ones. Yet, for all countries more sclerotic labor markets generate higher volatility than fluid ones conditional on discount shocks. The top panel of Table 2.6 shows that sclerotic markets are more important in France than in other countries in amplifying the variation in discounts.

We assess the role of LMIs on the propagation of productivity shocks with Figure 2.12. Consistently with the literature, our model with productivity shocks does not generate enough unemployment volatility to explain the data. Productivity does a worse job under sclerotic labor markets relative to fluid ones: as we observed in the right panel of Figure 2.8, productivity shocks cause more persistent but less volatile movements in unemployment. The bottom panel of Table 2.6 summarizes the results. The measure of "fit" for France slightly increases with the sclerotic calibration relative to the fluid one, but a closer inspection of the corresponding plot reveals that this is due to a better timing of the variations rather than to increased volatility. In Spain, productivity shocks cause the wrong signs in the variation of simulated unemployment. The joint dynamics of productivity and unemployment in Spain constitute a long-standing puzzle. As argued in Comin et al. (2019), this may be due to the reliance in Spain on fixed-term contracts in recent decades. During the Great Recession, workers in fixed-terms contracts, likely working

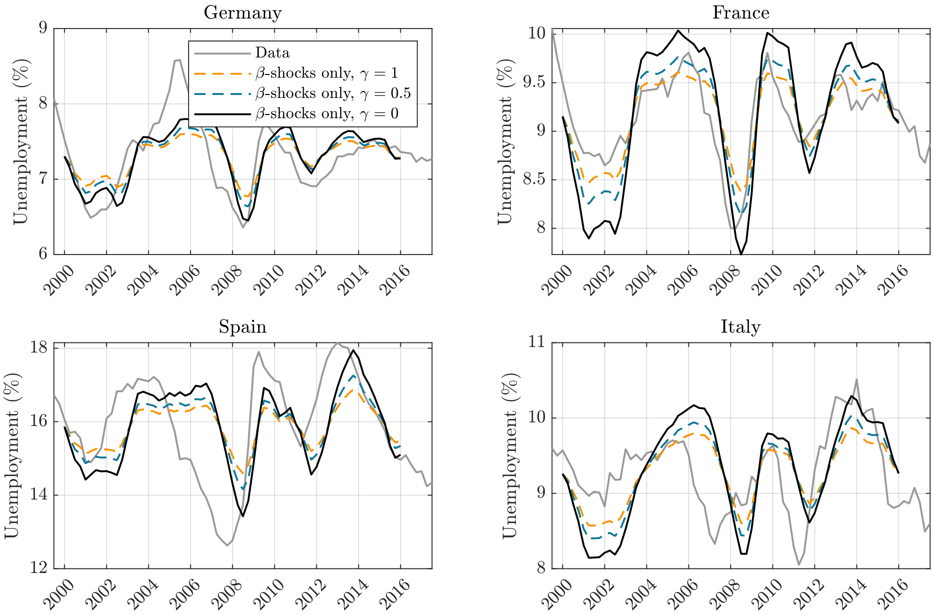

Figure 2.10: Simulated unemployment feeding in only SDF shocks, by degree of wage rigidity. Country specific calibration.

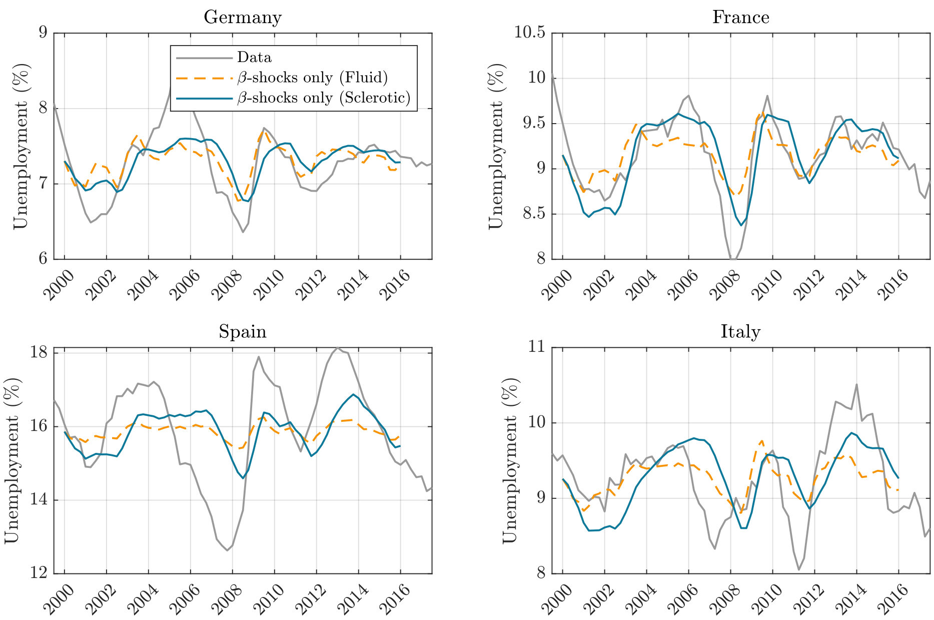

Figure 2.11: Simulated unemployment feeding in only SDF shocks, by fluidity of labor markets. Wages are fully flexible.

Table 2.6: Covariance between simulations and data (by LMI), relative to the volatility in observed (HP-filtered) unemployment. Fully flexible wages. Unemployment benefit set to $b = 0.4$.

LMIs	Germany	France	Spain	Italy
Only β-shocks				
Fluid	0.2841	0.4057	0.0889	0.2570
Sclerotic	0.3415	0.7627	0.1523	0.3087
Only z-shocks				
Fluid	0.0729	0.0422	−0.0947	0.1319
Sclerotic	0.0171	0.0592	−0.1181	0.0406

Table 2.7: Covariance between simulations and data, relative to the volatility in observed (HP-filtered) unemployment.

Source of variation	Germany	France	Spain	Italy
Only z-shocks	0.0344	0.1147	−0.2715	0.0776
Only β-shocks	0.3415	0.7627	0.1523	0.3087
Both shocks	0.3748	0.8781	−0.1191	0.3882

lower hours and at lower productivity than workers in fixed-term contracts, have been the first to lose employment. This may explain why both output per worker and unemployment have increased.

We conclude by comparing the relative contribution of SDF shocks to productivity shocks. Figure 2.13 shows simulated unemployment as predicted by both shocks fed in the model. Table 2.7 provides a numerical representation of the results. It also shows the simulations where one shock is shut down, in order to provide a sense of the decomposition of the overall effects. We see that the simulations with both shocks are predominantly driven by SDF shocks as opposed to productivity shocks. Quantitatively, our model fits France better than Germany and Italy, while we predict the wrong variations in Spain due to productivity.

2.5 Conclusion

In this paper we wrote a representative agent model of the labor market with search and matching frictions, augmented with a discount factor shock. We estimate discount factor and aggregate productivity shocks externally relative to the model. We simulated the model with the estimated shocks and compared the simulations to the data. We found that discount factor shocks can explain a significant portion of the volatility of unemployment, contrary to productivity shocks. However, neither the discount factor nor the productivity shocks can explain the differences in unemployment dynamics that we observe across European countries.

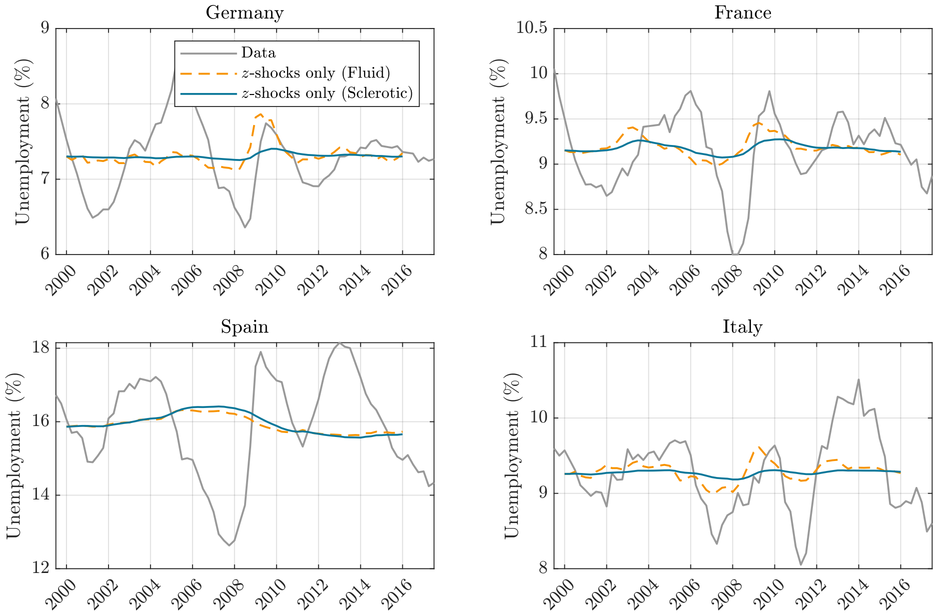

Figure 2.12: Simulated unemployment feeding in only productivity shocks, by fluidity of labor markets. Wages are fully flexible.

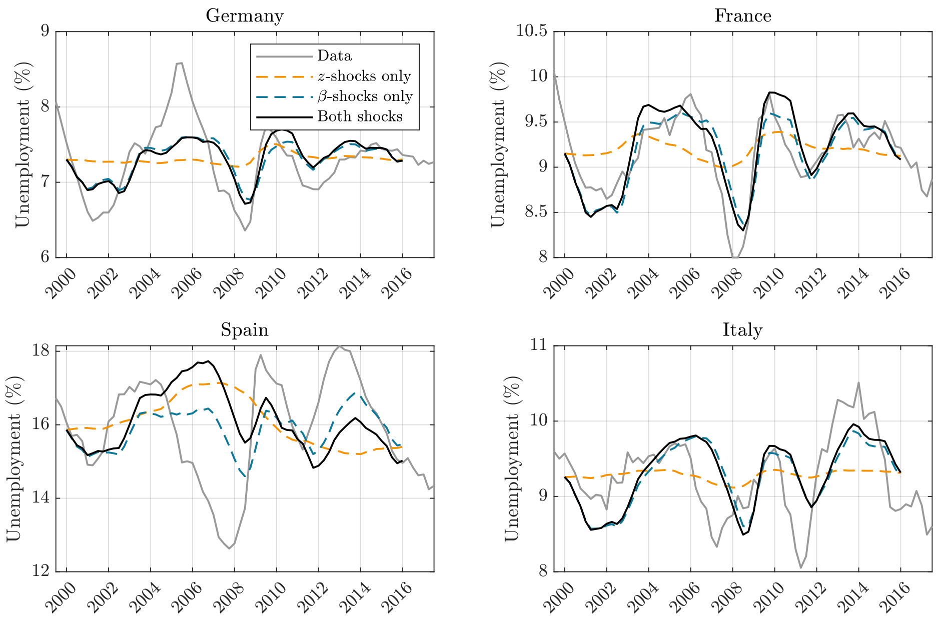

Figure 2.13: Decomposition of simulated unemployment feeding in both shocks. Wages are fully flexible. Country-specific calibration.

Chapter 3

European Unemployment and Dual Labor Markets

This chapter is based on joint work with Diego Comin, Riccardo Franceschin and Antonella Trigari.

3.1 Introduction

Aggregate data on labor market outcomes shows considerable differences across European countries. While the findings in the previous Chapter show that discounts can explain fluctuations in the time series of unemployment, discounts cannot explain the timing of certain variations. This paper explores the hypothesis that European countries are characterized by different uses of Fixed-Term Contracts (FTCs) and Open-Ended Contracts (OECs). An economy that predominantly uses OECs is plausibly adjusting to macroeconomic conditions more slowly relative to an economy that relies on FTCs. This Chapter presents the model and its main mechanisms.

We extend the Diamond-Mortensen-Pissarides (DMP) model in the previous Chapter by introducing FTCs and OECs. Their coexistence in equilibrium is motivated by heterogeneous match-specific workers' productivity. Firms and workers choose between the two contracts when matched. The FTC is more flexible relative to the OEC in that firms can terminate it without any firing cost. A FTC lasts only one period, but agents may choose to renew it. On the other hand, an OEC features a lower exogenous job-destruction rate relative to a FTC. We interpret the difference between the two exogenous job-destruction rates as different probabilities with which a worker quits a job for reasons not directly related to labor market outcomes.

Matches between firms and workers are heterogeneous because of idiosyncratic productivity, which is drawn at the match. However, matched agents initially observe only one signal about the match-specific productivity. They fully observe it only after some periods with some exogenous arrival rate.

Future work consists of replicating the methodology applied to the representative agent model with the heterogeneous agents, dual labor market model. The goal is to understand whether different implementations of this duality can account for differences in unemployment dynamics across EU countries.

We contribute to the dual labor market literature by providing a model of business cycles with a possible choice for agents between OEC and FTC. The choice that agents faces is similar to Garibaldi and Violante (2005), but with the addition of a learning process that is instead present in Faccini (2014). These aspects become fundamental in presence of shocks to productivity or to the discount factor, since different contracts can induce different responses of agents. FTCs for example are easy to terminate and they can be adjusted quickly increasing the unemployment level, as noted in Caggese and Cuñat (2008). Finally, these aspects seems relevant in determining the overall performance of a labor market after a shock, as noted in Bentolila et al. (2012).

Section 3.2 describes the model and its key mechanisms. Section 3.3 discusses the methodology and the empirical strategy, together with directions of future work.

3.2 The Model

3.2.1 Description of the model and main mechanisms

This model takes a standard Diamond-Mortensen-Pissarides model and adds a major departure: the presence of both temporary and permanent contracts. On one hand, temporary contracts can be endogenously destroyed at no cost, while permanent contracts can be destroyed subject to a firing cost. On the other hand, permanent contracts benefit from a lower probability of exogenous separation. To generate a trade-off between temporary and permanent contracts, we consider a distribution of match-specific productivities, which are first signalled and then randomly revealed at a later stage. When an unemployed worker and a firm are matched, a match-specific productivity a is drawn, but only a noisy signal s is observed by agents. Depending on the signal, agents decide to discard the match, sign a temporary contract or sign a permanent contract. After the first period of a contract, the productivity a is revealed with probability ξ through a Calvo lottery. Depending on the information available to them, agents decide whether to terminate an existing contract or to renew it. If a contract is temporary, then it can be converted to a permanent one. If a contract is permanent, then it can either be destroyed or renewed. However, if a contract is temporary, the option to renew it as such is unavailable to agents with probability ϕ, in which case they either destroy the contract or convert it into a permanent one.

Temporary contracts provide a way for agents to insure against the risk of low productivity a in face of a high signal s. Once such uncertainty is resolved, agents have no incentive to opt for a temporary contract other than the absence of a firing cost. The advantage of a permanent contract is the reduced risk of exogenous separation.

To understand better the model, let us consider a newly formed match and follow it through time. Matches that take place in period $t-1$ become productive only at period t. At the match, a signal s over the the match-specific productivity is known. At the beginning of period t, the pair (β_t, z_t) becomes common knowledge. As soon as aggregate uncertainty resolves, the worker and the firm involved in the new match bargain. Based on s and the aggregate state, firms and workers bargain on wages and they decide whether to reject the match, write a temporary contract or write a permanent contract. After the contract has been written, production happens. With some exogenous probabilities $\lambda^T > \lambda^P$, the match is broken: the worker will go back to unemployment and the firm will post a new vacancy in period $t+1$. If the match is not broken, then agents learn the true value of a with probability ξ. If they do not learn a, they retain their knowledge of s and $a|s$. Period t ends.

Period $t+1$ begins and the pair (β_{t+2}, z_{t+1}) becomes common knowledge. Based on either s or a and the aggregate state, the worker and the firm bargain on the wage for period $t+1$ and decide whether to separate or continue the contract. If a contract continues, then it can be again either temporary or permanent. However, with probability ϕ, the option of *keeping* a temporary contract is unavailable to agents. As the wage at $t+1$ is bargained *before* the new contract is written and *before* production takes place, wages for the second period of a contract are going to depend on the type of contract that was in place in period t. In particular, the wage for a permanent contract at $t+1$ that was temporary at t will differ from the wage for a permanent contract at $t+1$ that was permanent at t. This happens because the firm has different outside options. If the first contract was temporary, then the firm may decide to fire the worker at no cost. If the first contract was permanent, then the firm may decide to fire the worker at a fixed firing cost f. After the contract for period $t+1$ has been written production takes place. With some exogenous probability λ^T or λ^P, the match is broken. If agents did not learn a at the end of the previous period, then they will have the chance to learn it with probability ξ. Period $t+1$ ends.

If a match survives this far, then the dynamics of period $t+1$ will repeat in subsequent periods.

3.2.2 Timing

Within each period t, the following happens, in this order.

1. Aggregate uncertainty (β_t, z_t) resolves.

2. New matches are formed: both a and s are drawn, but only s is observed.

3. All agents in a match (newly formed or not) bargain over wages and contracts on the basis of s or a depending on their information set. Agents in an old match that are bound by a temporary contract are kept from renewing it as temporary with probability ϕ (and are so left with either endogenously breaking the contract or transorming it to a permanent).

4. Production takes place.

5. Matches are exogenously destroyed with probability λ^T or λ^P, depending on which contract is in place.

6. Agents that know s but not a gain knowledge of a with probability ξ. Those who do not learn a retain their knowledge of s and $a|s$.

3.2.3 The model

The choice of endogenously terminating a contract is a bilateral decision from both the firm and the worker. In other words, if the firm finds it convenient to fire the worker, also the worker will find it convenient to return to unemployment. This assumption simplifies the exposition and the solution of the model. We refer to the endogenous separations as the firm deciding to fire the worker. We interpret exogeonus separations as matches that break for reasons we cannot capture in the model (e.g., a firm ceasing its activity for reasons not related to the labor market). In case of an exogenous separation, the firm never incurs in any firing cost.

Idiosyncratic (match-specific) productivity

$$a \overset{iid}{\sim} \mathcal{N}(\mu_a, \sigma_a)$$

Signal to employer at the match

$$s = a + \sigma_s \varepsilon_t^s \qquad \varepsilon_t^s \overset{iid}{\sim} \mathcal{N}(0, 1)$$

The prior is $a \sim \mathcal{N}(\mu_a, \sigma_a^2)$, likelihood is $s|a \sim \mathcal{N}(a, \sigma_s^2)$. The posterior is

$$a|s \sim \mathcal{N}\left(\frac{s\sigma_a^2 + \mu_a\sigma_s^2}{\sigma_a^2 + \sigma_s^2}, \frac{\sigma_a^2\sigma_s^2}{\sigma_a^2 + \sigma_s^2}\right).$$

Let $F_{a|s}(\cdot)$ denote the CDF of $a|s$.

The exogenous processes are the aggregate discount factor β and the aggregate productivity z. They evolve according to AR(1) processes:

$$z_t = (1-\rho_z)\bar{z} + \rho_z z_{t-1} + \sigma_z \varepsilon_t^z \qquad \varepsilon_t^z \overset{iid}{\sim} \mathcal{N}(0,1)$$
$$\beta_t = (1-\rho_\beta)\bar{\beta} + \rho_\beta \beta_{t-1} + \sigma_\beta \varepsilon_t^\beta \qquad \varepsilon_t^\beta \overset{iid}{\sim} \mathcal{N}(0,1)$$

Note that β_t is subject to the shock ε_t^β. In other words, the value β_t is known at the beginning of period t.

In the following, the expectation operator $\mathbf{E}_t(\cdot)$ is taken with respect to *aggregate* uncertainty, which is here given by productivity z_t and discounts β_t. The expectation with respect to idiosyncratic uncertainty is explicitly written with an integral.

As anticipated above, we have two relevant categories of periods. The first one is about the first periods of a contract, from the match up until the learning of the match-specific productivity (unless the match is broken before this moment). In these periods, agents take decisions based on their knowledge of s and the aggregate state. Each contract can be either permanent or temporary. The values and the wages in this category of periods present the superscript T or P to reflect the choice of the contract. After the first period of a match, the type of the contract in the previous period is relevant for wage bargaining as mentioned above. We keep track of this by attaching the superscripts $\{T,T\}$, $\{T,P\}$ and $\{P,P\}$ to the value functions and the wages. Values and wages in this category of periods are function of s and not of a.

The second type of periods is about the later periods of a contract starting from the moment where agents learn about a. Again, each contract can be either temporary or permanent and the type of the previous contract is relevant. Values and wages in this category of periods are function of a and not of s. We group the exposition of the value functions, surpluses and wages by these periods.

Periods of a match where only s is known

Let J_t^i denote the value of a temporary ($i = T$) or a permanent ($i = P$) job at the first period after a match. Let the value of a worker W_t be superscripted similarly.

At the match, a match-specific productivity a is drawn. However, at this stage workers and firms only observe the noisy signal $s = a + \varepsilon$. At the match, workers and firms bargain a wage and a contract type given their knowledge of s. Agents can decide to reject the match if the signal s is too low. Otherwise they decide whether to start a temporary contract or a permanent one. The firm has the following values for a job at this stage and at this first period of a match, $J_t^T(s)$ and $J_t^P(s)$:

$$\begin{aligned} J_t^T(s) = &\int_{-\infty}^{\infty} a \, \mathrm{d}F_{a|s}(a) + z_t - w_t^T(s) + \beta_t \mathbf{E}_t \Bigg(\lambda^T V_{t+1} + (1-\lambda^T)\xi \times \\ &\times \left[(1-\phi) \int_{-\infty}^{\infty} \max\left\{ V_{t+1}; J_{t+1}^{T,T}(a); J_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + \phi \int_{-\infty}^{\infty} \max\left\{ V_{t+1}; J_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) \right] + \\ &+ (1-\lambda^T)(1-\xi) \left[(1-\phi) \max\left\{ V_{t+1}; J_{t+1}^{T,T}(s); J_{t+1}^{T,P}(s) \right\} + \phi \max\left\{ V_{t+1}; J_{t+1}^{T,P}(s) \right\} \right] \Bigg) \\ J_t^P(s) = &\int_{-\infty}^{\infty} a \, \mathrm{d}F_{a|s}(a) + z_t - w_t^P(s) + \beta_t \mathbf{E}_t \Bigg(\lambda^P V_{t+1} + (1-\lambda^P) \times \\ &\times \left[\xi \int_{-\infty}^{\infty} \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + (1-\xi) \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(s) \right\} \right] \Bigg) . \end{aligned}$$

The firm observes the signal s and computes the expected match-specific productivity given the posterior distribution of a, $F_{a|s}(a)$. It also collects the aggregate productivity z_t. It pays the wage, which depends on the type of contract. With probability λ^T (λ^P), the temporary (permanent) contract is exogenously destroyed and the firm posts a new vacancy. If the match is not exogenously destroyed, then the firm decides whether to endogenously fire the worker and collect V or to keep the contract. If the existing contract is temporary, it can either be renewed as such or converted to a permanent contract. The choice of keeping the temporary contract is unavailable to agents with probability ϕ. If the existing contract is permanent, the firm will have to pay a fixed firing cost f in order to fire the worker. With probability ξ, agents learn the true match-specific productivity a. If they do not learn it, then they retain the knowledge of s as drawn at the beginning of the match, together with the posterior distribution $a|s$.

On the other hand the worker collects the values of working $W_t^T(s)$ and $W_t^P(s)$:

$$
\begin{aligned}
W_t^T(s) = w_t^T(s) + \beta_t \mathbf{E}_t \Bigg(& \lambda^T U_{t+1} + (1-\lambda^T)\xi \left[(1-\phi) \int_{-\infty}^{\infty} \max\left\{ U_{t+1}; W_{t+1}^{T,T}(a); W_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + \right. \\
& \left. +\phi \int_{-\infty}^{\infty} \max\left\{ U_{t+1}; W_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) \right] + \\
& + (1-\lambda^T)(1-\xi) \left[(1-\phi) \max\left\{ U_{t+1}; W_{t+1}^{T,T}(s); W_{t+1}^{T,P}(s) \right\} + \phi \max\left\{ U_{t+1}; W_{t+1}^{T,P}(s) \right\} \right] \Bigg) \\
W_t^P(s) = w_t^P(s) + \beta_t \mathbf{E}_t \Bigg(& \lambda^P U_{t+1} + (1-\lambda^P) \left[\xi \int_{-\infty}^{\infty} \max\left\{ U_{t+1}; W_{t+1}^{P,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + \right. \\
& \left. +(1-\xi) \max\left\{ U_{t+1}; W_{t+1}^{P,P}(s) \right\} \right] \Bigg).
\end{aligned}
$$

Given the signal, the worker collects the wage, which again differs according to the type of contract. If the match is exogenously destroyed with probability λ^T or λ^P, the worker collects the expected value of unemployment. If the match is not exogenously destroyed, then the worker either continues working or goes back to unemployment if the firm fires her.

After the first period of the match, agents may still ignore the true value of a. In this case, the value functions for the firm are

$$
\begin{aligned}
J_t^{T,T}(s) = \int_{-\infty}^{\infty} & a \, \mathrm{d}F_{a|s}(a) + z_t - w_t^{T,T}(s) + \beta_t \mathbf{E}_t \Bigg(\lambda^T V_{t+1} + (1-\lambda^T)\xi \left[(1-\phi) \times \right. \\
& \left. \times \int_{-\infty}^{\infty} \max\left\{ V_{t+1}; J_{t+1}^{T,T}(a); J_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + \phi \int_{-\infty}^{\infty} \max\left\{ V_{t+1}; J_{t+1}^{T,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) \right] + \\
& + (1-\lambda^T)(1-\xi) \left[(1-\phi) \max\left\{ V_{t+1}; J_{t+1}^{T,T}(s); J_{t+1}^{T,P}(s) \right\} + \phi \max\left\{ V_{t+1}; J_{t+1}^{T,P}(s) \right\} \right] \Bigg) \\
J_t^{T,P}(s) = \int_{-\infty}^{\infty} & a \, \mathrm{d}F_{a|s}(a) + z_t - w_t^{T,P}(s) + \beta_t \mathbf{E}_t \Bigg(\lambda^P V_{t+1} + (1-\lambda^P)\xi \times \\
& \times \int_{-\infty}^{\infty} \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + (1-\lambda^P)(1-\xi) \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(s) \right\} \Bigg) \\
J_t^{P,P}(s) = \int_{-\infty}^{\infty} & a \, \mathrm{d}F_{a|s}(a) + z_t - w_t^{P,P}(s) + \beta_t \mathbf{E}_t \Bigg(\lambda^P V_{t+1} + (1-\lambda^P)\xi \times \\
& \times \int_{-\infty}^{\infty} \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(a) \right\} \, \mathrm{d}F_{a|s}(a) + (1-\lambda^P)(1-\xi) \max\left\{ V_{t+1} - f; J_{t+1}^{P,P}(s) \right\} \Bigg)
\end{aligned}
$$

and for the worker are

$$
\begin{aligned}
W_t^{T,T}(s) = w_t^{T,T}(s) + \beta_t \mathbf{E}_t \Bigg(& \lambda^T U_{t+1} + (1-\lambda^T)\xi \Bigg[(1-\phi)\int_{-\infty}^{\infty} \max\left\{U_{t+1}; W_{t+1}^{T,T}(a); W_{t+1}^{T,P}(a)\right\} \, \mathrm{d}F_{a|s}(a) + \\
& +\phi \int_{-\infty}^{\infty} \max\left\{U_{t+1}; W_{t+1}^{T,P}(a)\right\} \, \mathrm{d}F_{a|s}(a)\Bigg] + \\
& + (1-\lambda^T)(1-\xi)\Big[(1-\phi)\max\left\{U_{t+1}; W_{t+1}^{T,T}(s); W_{t+1}^{T,P}(s)\right\} + \phi \max\left\{U_{t+1}; W_{t+1}^{T,P}(s)\right\}\Big]\Bigg) \\
W_t^{T,P}(s) = w_t^{T,P}(s) + \beta_t \mathbf{E}_t \Bigg(& \lambda^P U_{t+1} + (1-\lambda^P)\Bigg[\xi \int_{-\infty}^{\infty} \max\left\{U_{t+1}; W_{t+1}^{P,P}(a)\right\} \, \mathrm{d}F_{a|s}(a) + \\
& + (1-\xi)\max\left\{U_{t+1}; W_{t+1}^{P,P}(s)\right\}\Bigg]\Bigg) \\
W_t^{P,P}(s) = w_t^{P,P}(s) + \beta_t \mathbf{E}_t \Bigg(& \lambda^P U_{t+1} + (1-\lambda^P)\Bigg[\xi \int_{-\infty}^{\infty} \max\left\{U_{t+1}; W_{t+1}^{P,P}(a)\right\} \, \mathrm{d}F_{a|s}(a) + \\
& + (1-\xi)\max\left\{U_{t+1}; W_{t+1}^{P,P}(s)\right\}\Bigg]\Bigg).
\end{aligned}
$$

The meaning of these values is analogous to the Bellman equations above. The main difference lies in the wages, which depend on the type of the previous contract because of different outside options for the firm.

The surpluses from temporary and permanent contracts *before such contracts are signed* are

$$
\begin{aligned}
S_t^T(s) &= \left[J_t^T(s) - V_t\right] + \left[W_t^T(s) - U_t\right] \\
S_t^P(s) &= \left[J_t^P(s) - V_t\right] + \left[W_t^P(s) - U_t\right] \\
S_t^{T,T}(s) &= \left[J_t^{T,T}(s) - V_t\right] + \left[W_t^{T,T}(s) - U_t\right] \\
S_t^{T,P}(s) &= \left[J_t^{T,P}(s) - V_t\right] + \left[W_t^{T,P}(s) - U_t\right] \\
S_t^{P,P}(s) &= \left[J_t^{P,P}(s) - V_t\right] + \left[W_t^{P,P}(s) - U_t\right]
\end{aligned}
$$

The surpluses of the firm $J_t^P(s)$ and $J_t^{T,P}$ do not include the firing cost because the surpluses are measured *before* the contract in period t is signed, so that the firm may simply decide not to sign a contract if the permanent one is not convenient.

Wages are set at the match according to Nash bargaining.

$$
\begin{aligned}
w_t^T(s) &= \arg\max_{w_t^T(s)} \left[J_t^T(s) - V_t\right]^{\eta}\left[W_t^T(s) - U_t\right]^{1-\eta} \\
w_t^P(s) &= \arg\max_{w_t^P(s)} \left[J_t^P(s) - V_t\right]^{\eta}\left[W_t^P(s) - U_t\right]^{1-\eta} \\
w_t^{T,T}(s) &= \arg\max_{w_t^{T,T}(s)} \left[J_t^{T,T}(s) - V_t\right]^{\eta}\left[W_t^{T,T}(s) - U_t\right]^{1-\eta} \\
w_t^{T,P}(s) &= \arg\max_{w_t^{T,P}(s)} \left[J_t^{T,P}(s) - V_t\right]^{\eta}\left[W_t^{T,P}(s) - U_t\right]^{1-\eta} \\
w_t^{P,P}(s) &= \arg\max_{w_t^{P,P}(s)} \left[J_t^{P,P}(s) - (V_t - f)\right]^{\eta}\left[W_t^{P,P}(s) - U_t\right]^{1-\eta}.
\end{aligned}
$$

Note that the relevant surplus of the firm in determining the wage of a permanent contract, $w_t^P(s)$, is simply $J_t^P(s) - V_t$ and not $J_t^P(s) - V_t - f$. The argument is analogous in the case of $w_t^{T,P}(s)$.

Periods of a match where a is known

Aggregate uncertainty resolves at the beginning of the period. Before production, the parties renegotiate the wage on the basis of a. Contracts that were previously temporary may be kept temporary or converted to permanent, or they can be rescinded. The option of keeping a temporary contract is unavailable to agents with probability ϕ. Contracts that were permanent can only be kept permanent or rescinded. If the previous contract was temporary, and because such contract is in place at the moment of the renegotiation, the firm does not consider the firing cost when bargaining the wage for the second period. For this reason we need to keep track of the contract type in the previous period.

The firm's value of a job are the following:

$$\begin{aligned} J_t^{T,T}(a) &= a + z_t - w_t^{T,T}(a) + \beta_t \mathbf{E}_t \left(\lambda^T V_{t+1} + (1-\lambda^T)\left[(1-\phi)\max\left\{V_{t+1}; J_{t+1}^{T,T}(a); J_{t+1}^{T,P}(a)\right\} + \right.\right. \\ &\qquad \left.\left. +\phi \max\left\{V_{t+1}; J_{t+1}^{T,P}(a)\right\}\right]\right) \\ J_t^{T,P}(a) &= a + z_t - w_t^{T,P}(a) + \beta_t \mathbf{E}_t \left(\lambda^P V_{t+1} + (1-\lambda^P)\max\left\{V_{t+1} - f; J_{t+1}^{P,P}(a)\right\}\right) \\ J_t^{P,P}(a) &= a + z_t - w_t^{P,P}(a) + \beta_t \mathbf{E}_t \left(\lambda^P V_{t+1} + (1-\lambda^P)\max\left\{V_{t+1} - f; J_{t+1}^{P,P}(a)\right\}\right). \end{aligned}$$

Firms collect production $a + z_t$ and pay wages. If the contract is not exogenously destroyed with probability λ^P, then the firm decides whether to keep the worker or to fire her.

Workers' value of working is similar as above, with the exception of the wage.

$$\begin{aligned} W_t^{T,T}(a) &= w_t^{T,T}(a) + \beta_t \mathbf{E}_t \left(\lambda^T U_{t+1} + (1-\lambda^T)\left[(1-\phi)\max\left\{U_{t+1}; W_{t+1}^{T,T}(a); W_{t+1}^{T,P}(a)\right\} + \right.\right. \\ &\qquad \left.\left. +\phi \max\left\{U_{t+1}; W_{t+1}^{T,P}(a)\right\}\right]\right) \\ W_t^{T,P}(a) &= w_t^{T,P}(a) + \beta_t \mathbf{E}_t \left(\lambda^P U_{t+1} + (1-\lambda^P)\max\left\{U_{t+1}; W_{t+1}^{P,P}(a)\right\}\right) \\ W_t^{P,P}(a) &= w_t^{P,P}(a) + \beta_t \mathbf{E}_t \left(\lambda^P U_{t+1} + (1-\lambda^P)\max\left\{U_{t+1}; W_{t+1}^{P,P}(a)\right\}\right). \end{aligned}$$

Surpluses after the first period of a contract *before* the new temporary/permanent contract is signed are

$$\begin{aligned} S_t^{T,T}(a) &= \left[J_t^{T,T}(a) - V_t\right] + \left[W_t^{T,T}(a) - U_t\right] \\ S_t^{T,P}(a) &= \left[J_t^{T,P}(a) - V_t\right] + \left[W_t^{T,P}(a) - U_t\right] \\ S_t^{P,P}(a) &= \left[J_t^{P,P}(a) - (V_t - f)\right] + \left[W_t^{P,P}(a) - U_t\right]. \end{aligned}$$

Wages in this period are again Nash-bargained:

$$\begin{aligned} w_t^{T,T}(a) &= \arg\max_{w_t^{T,T}(a)} \left[J_t^{T,T}(a) - V_t\right]^{\eta}\left[W_t^{T,T}(a) - U_t\right]^{1-\eta} \\ w_t^{T,P}(a) &= \arg\max_{w_t^{T,P}(a)} \left[J_t^{T,P}(a) - V_t\right]^{\eta}\left[W_t^{T,P}(a) - U_t\right]^{1-\eta} \\ w_t^{P,P}(a) &= \arg\max_{w_t^{P,P}(a)} \left[J_t^{P,P}(a) - (V_t - f)\right]^{\eta}\left[W_t^{P,P}(a) - U_t\right]^{1-\eta} \end{aligned}$$

Again, the firm does not consider firing costs if the previous contract was temporary, as this renegotiation happens with the previous contract in place.

Other equations of the model

Let $F_s(\cdot)$ denote the CDF of the signals s. The value of unemployment for a worker is:

$$U_t = b + \beta_t \mathbf{E}_t \left((1-p_t)U_{t+1} + p_t \int_{-\infty}^{\infty} \max\left\{U_{t+1}; W_{t+1}^T(s); W_{t+1}^P(s)\right\} \, \mathrm{d}F_s(s)\right).$$

The unemployed worker collects the unemployment benefit b. She finds a job with probability p_t If she finds a job, then she will complete the unemployment spell for period t and the match will be effective at time $t+1$, at which point she may either get a temporary contract or a permanent one depending on the (unknown at this stage) draw of the signal s. If she does not find a job, she continues collecting the value of unemployment.

Firms need to post vacancies before hiring. The value of opening a vacancy is

$$V_t = -\kappa + \beta_t \mathbf{E}_t \left((1-q_t)V_{t+1} + q_t \int_{-\infty}^{\infty} \max\left\{ V_{t+1}; J_{t+1}^T(s); J_{t+1}^P(s) \right\} \ \mathrm{d}F_s(s) \right).$$

The firm faces a fixed cost κ to open a vacancy. It will find a worker with probability q_t, but the match is assumed to be effective starting in period $t+1$. In this case, the firm will decide whether to sign a temporary or permanent contract on the basis of the (unknown at this stage) signal s. If the firm does not find a worker, then it collects the future value of a vacancy.

We assume free-entry for firms, which drives the value of a vacancy to zero. By setting $V_t = 0$ at all periods t, we have

$$\frac{\kappa}{q_t} = \beta_t \mathbf{E}_t \left(\int_{-\infty}^{\infty} \max\left\{ 0; J_{t+1}^T(s); J_{t+1}^P(s) \right\} \ \mathrm{d}F_s(s) \right).$$

Matches depend on the stock of unemployed people and the number of open vacancies:

$$m_t = \sigma_m u_t^{\sigma} v_t^{1-\sigma}.$$

The transition probabilities are:

$$p_t = \frac{m_t}{u_t} \qquad\qquad q_t = \frac{m_t}{v_t}.$$

Let $C_t(\cdot)$ denote the contract that is chosen by firms and workers. Formally,

$$\begin{aligned}
C_t^0(s) &\equiv \arg\max\left\{ V_t; J_t^T(s); J_t^P(s) \right\} \in \{N, T, P\} \\
C_t^T(s) &\equiv \arg\max\left\{ V_t; J_t^{T,T}(s); J_t^{T,P}(s) \right\} \in \{N, T, P\} \\
C_t^P(s) &\equiv \arg\max\left\{ V_t - f; J_t^{P,P}(s) \right\} \in \{N, P\} \\
C_t^{\phi}(s) &\equiv \arg\max\left\{ V_t; J_t^{T,P}(s) \right\} \in \{N, P\} \\
C_t^T(a) &\equiv \arg\max\left\{ V_t; J_t^{T,T}(a); J_t^{T,P}(a) \right\} \in \{N, T, P\} \\
C_t^P(a) &\equiv \arg\max\left\{ V_t - f; J_t^{P,P}(a) \right\} \in \{N, P\} \\
C_t^{\phi}(a) &\equiv \arg\max\left\{ V_t; J_t^{T,P}(a) \right\} \in \{N, P\},
\end{aligned}$$

where N stands for none, T for temporary and P for permanent.

With these definitions, we start the following accounting exercise. The probabilities to sign no/a temporary/a permanent contract in the first period of the matches are

$$\begin{aligned}
N_t^0 &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^0(s) = N \right) \ \mathrm{d}F_s(s) \\
T_t^0 &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^0(s) = T \right) \ \mathrm{d}F_s(s) \\
P_t^0 &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^0(s) = P \right) \ \mathrm{d}F_s(s).
\end{aligned}$$

In words, N_t^0 is the probability that new matches that are rejected, while T_t^0 and P_t^0 are the probabilities that temporary and permanent contracts are signed right after a match. The probabilities for contracts after the first period of a match that were previously temporary contracts are

$$\begin{aligned} N_t^{T;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(s) = N\right)\, \mathrm{d}F_s(s) & N_t^{T;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(a) = N\right)\, \mathrm{d}F_a(a) \\ T_t^{T;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(s) = T\right)\, \mathrm{d}F_s(s) & T_t^{T;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(a) = T\right)\, \mathrm{d}F_a(a) \\ P_t^{T;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(s) = P\right)\, \mathrm{d}F_s(s) & P_t^{T;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^T(a) = P\right)\, \mathrm{d}F_a(a), \end{aligned}$$

those that were previously permanent are

$$\begin{aligned} N_t^{P;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^P(s) = N\right)\, \mathrm{d}F_s(s) & N_t^{P;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^P(a) = N\right)\, \mathrm{d}F_a(a) \\ P_t^{P;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^P(s) = P\right)\, \mathrm{d}F_s(s) & P_t^{P;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^P(a) = P\right)\, \mathrm{d}F_a(a), \end{aligned}$$

and those that were temporary and cannot be renewed (because of the lottery with probability ϕ)

$$\begin{aligned} N_t^{\phi;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^\phi(s) = N\right)\, \mathrm{d}F_s(s) & N_t^{\phi;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^\phi(a) = N\right)\, \mathrm{d}F_a(a) \\ P_t^{\phi;s} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^\phi(s) = P\right)\, \mathrm{d}F_s(s) & P_t^{\phi;a} &\equiv \int_{-\infty}^{\infty} \mathbf{1}\left(C_t^\phi(a) = P\right)\, \mathrm{d}F_a(a). \end{aligned}$$

In words (and taking an example), $P_t^{T,a}$ is the probability that a worker signs a permanent contract after coming from a temporary and given that she observed the match-specific productivity a.

We measure the stocks of workers that are employed with a temporary/permanent contract depending on whether they only know s or they already know a. The mass of workers that obtain a job is:

$$\begin{aligned} e_{t+1}^{T;s} &= p_t u_t T_t^0 \\ e_{t+1}^{P;s} &= p_t u_t P_t^0, \end{aligned}$$

In essence, this is considering workers that find a job from unemployment ($p_t u_t$) and that sign a temporary/permanent contract. There cannot be stocks $e_t^{\cdot;a}$ because nobody learns a right after the match. The stocks of workers that are employed after the first period of a contract given they were employed under a temporary/permanent are:

$$\begin{aligned} e_{t+1}^{T,T} &= e_t^{T,T} + e_t^T(1-\lambda^T)T_t^T - e_t^{T,T}\left[\lambda^T + (1-\lambda^T)\left[\phi N_t^\phi + (1-\phi)N_t^T\right]\right] \\ e_{t+1}^{T,P} &= e_t^T(1-\lambda^T)P_t^T \\ e_{t+1}^{P,P} &= e_t^{P,P} + e_t^P(1-\lambda^P)P_t^P - e_t^{P,P}\left[\lambda^P + (1-\lambda^P)N_t^P\right]. \end{aligned}$$

Each worker transitions to a permanent contract provided that the match is not exogenously broken and provided that both the firm and the worker found it profitable to sign the new permanent contract. The transitions from permanent to permanent need to account for workers whose contract lasted for three periods or more, together with those that will become unemployed for either exogenous or endogenous reasons. Note that $e_t^{T,P}$ is reset at every period (i.e., it does not depend on its past value).

Unemployment evolves according to the following law of motion

$$\begin{aligned} u_{t+1} = u_t &- m_t\left[T_t^0 + P_t^0\right] + \lambda^T\left[e_t^T + e_t^{T,T}\right] + \lambda^P\left[e_t^P + e_t^{T,P} + e_t^{P,P}\right] + \\ &+ (1-\lambda^T)N_t^T\left[e_t^T + e_t^{T,T}\right] + (1-\lambda^P)N_t^P\left[e_t^P + e_t^{T,P} + e_t^{P,P}\right]. \end{aligned}$$

3.3 Discussion

This paper only presents a model and its main mechanisms, together with some motivating observations. Because the calibration is not ready yet, we are unable to present Impulse-Response Functions (IRFs) nor simulations. We also cannot explore the necessary conditions of the model, because heterogeneity constitutes considerable complexity in the analytical tractability of the model. However, we can discuss certain properties of the model and outline a roadmap for future work.

First, because all value functions are increasing in the match-specific productivity, there are two endogenous thresholds. Consider a worker-firm pair that just got matched and who observe the productivity signal s. Let $\underline{s}$ and $\overline{s}$ be two endogenous thresholds, such that $\underline{s} < \overline{s}$. If $s < \underline{s}$, then the match will be jointly rejected by the parties. We interpret this as the parties not finding the match good enough. For example, the worker meets their potential colleagues, sees that their professional relationship would hinder the worker's productivity and would rather not work at that firm. If $\underline{s} < s < \overline{s}$, then the parties will jointly decide to sign a fixed-term contract. We interpret this as parties "trying out" the professional relationship before committing to a longer-term contract. This can either happen for workers that just entered the labor force (e.g., recent graduates) or for workers that change industry. If $s > \overline{s}$, then the parties will jointly decide to sign an open-ended contract. This may happen because parties are confident enough about the outcomes of the match and would rather secure the match for a high enough number of periods. Similar thresholds exist for the actual productivity a, which will be relevant when parties learn the true underlying match productivity. Therefore, we can partition the entire labor force in three segments: unemployed, employed with a FTC and employed with an OEC. The measures of these segments will be related to the thresholds. The thresholds $\underline{s}$, $\overline{s}$, $\underline{a}$ and $\overline{a}$ in a stationary steady-state will entirely depend on the calibration. This will allow us to target the fractions of people in each segment of the labor force that we observe in the data. However, doing so is difficult, because the distribution of wages will depends on these thresholds as well. This difficulty holds us back from disclosing the details of the calibration at this stage.

Second, agents in this model may wish to sign a FTC exclusively because it provides a probationary period. In fact, signing an OEC while the match-specific productivity is only noisily observed may be non-optimal. As soon as the real productivity is observed, agents have little to no incentive to keep a FTC. On the other hand, agents may spend a long period of time in a FTC because they take a long time to learn the real productivity. This may create a counterfactual prediction. In reality, many legislations in Europe require that a FTC does not last too long. While FTCs may be renewed a limited number of times, they are supposed to either be withdrawn or converted in an OEC. Therefore, we introduced a Calvo-style lottery whereby, with some probability, an existing FTC cannot be renewed as such. We will calibrate such probability to match the maximum duration a (renewed) FTC may have given existing legislations.

Third, the model aims at explaining aggregate phenomena, although it features heterogeneous agents. As mentioned above, they key aggregate moments we will target with the calibration are the fractions of workers that are unemployed, employed with a FTC and unemployed with an OEC. Additionally, we will give attention to labor market transitions across the three states. Particular attention will be given to the distribution of wages across employment statuses. This requires data on wages that is not necessarily available for all countries we focused on in the previous Chapter (i.e., Germany, France, Spain and Italy). Currently, we only have all the required data for Italy. Future work will investigate the availability of data for other countries.

Results we are currently working on are the following. First, as mentioned, we are working on the calibration of the model. Second, we will study the Impulse-Response Functions predicted by the model. This will allow us to inspect the main mechanisms a describe all the properties of the model. Third, we intend to provide simulations of the model given the exogenous series for productivity and discounts we obtained in the previous Chapter. While we established that productivity and discounts, together with wage rigidity, may explain the volatility of unemployment, adding dual labor markets will shed light on the differences in the timings of fluctuations.

Bibliography

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451.

Albertini, J. and Poirier, A. (2014). Discount Factor Shocks and Labor Market Dynamics. SFB 649 Discussion Paper 2014-033.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4):589–609.

Anginer, D., Demirguc-Kunt, A., and Zhu, M. (2014). How Does Competition Affect Bank Systemic Risk? *Journal of Financial Intermediation*, 23(1):1 – 26.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2006). Bank Concentration, Competition, and Crises: First Results. *Journal of Banking & Finance*, 30(5):1581 – 1603.

Bentolila, S., Cahuc, P., Dolado, J. J., and Le Barbanchon, T. (2012). Two-tier Labour Markets in the Great Recession: France Versus Spain. *The Economic Journal*, 122(562):F155–F187.

Berger, A. N., Klapper, L. F., and Turk-Ariss, R. (2009). Bank Competition and Financial Stability. *Journal of Financial Services Research*, 35:99–118.

Blanchard, O. J. and Diamond, P. (1989). The Aggregate Matching Function. Technical report, National Bureau of Economic Research, Cambridge, MA.

Boone, J. (2008). A New Way to Measure Competition. *The Economic Journal*, 118(531):1245–1261.

Boone, J., Griffith, R., and Harrison, R. (2005). Measuring Competition. Advanced Institute of Management Research Paper No. 022.

Borovička, J. and Borovičková, K. (2018). Risk Premia and Unemployment Fluctuations. New York University.

Boyd, J. H. and De Nicoló, G. (2005). The theory of bank risk taking and competition revisited. *The Journal of Finance*, 60(3):1329–1343.

Caggese, A. and Cuñat, V. (2008). Financing Constraints and Fixed-Term Employment Contracts. *The Economic Journal*, 118(533):2013–2046.

Campbell, J. Y. and Shiller, R. J. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *The Review of Financial Studies*, 1(3):195–228.

Chodorow-Reich, G. and Karabarbounis, L. (2016). The Cyclicality of the Opportunity Cost of Employment. *Journal of Political Economy*, 124(6):1563–1618.

Comin, D., Quintana-Gonzales, J., Schmitz, T. G., and Trigari, A. (2019). A New Measure of Utilization-Adjusted Total Factor Productivity Growth for European Countries.

De Loecker, J. and Eeckhout, J. (2019). The Rise of Market Power and the Macroeconomic Implications.

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–71.

Elsby, M. W. L., Hobijn, B., and Şahin, A. (2013). Unemployment Dynamics in the OECD. *Review of Economics and Statistics*, 95(2):530–548.

Faccini, R. (2014). Reassessing Labour Market Reforms: Temporary Contracts as a Screening Device. *The Economic Journal*, 124(575):167–200.

Garibaldi, P. and Violante, G. L. (2005). The Employment Effects of Severance Payments with Wage Rigidities. *The Economic Journal*, 115(506):799–832.

Hagedorn, M. and Manovskii, I. (2008). The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, 98(4):1692–1706.

Hall, R. E. (1988). The Relation between Price and Marginal Cost in U.S. Industry. *Journal of Political Economy*, 96(5):921–947.

Hall, R. E. (2017). High Discounts And High Unemployment. *American Economic Review*, 107(2):305–330.

Hellmann, T. F., Murdock, K. C., and Stiglitz, J. E. (2000). Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough? *American Economic Review*, 90(1):147–165.

Hillegeist, S. A., Keating, E. K., Cram, D. P., and Lundstedt, K. G. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9:5–34.

Hosios, A. J. (1990). On the Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies*, 57(2):279.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2):449–470.

Morlacco, M. (2019). Market Power in Input Markets: Theory and Evidence from French Manufacturing.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1):109–131.

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297.

Panzar, J. C. and Rosse, J. N. (1987). Testing For "Monopoly" Equilibrium. *The Journal of Industrial Economics*, 35(4):443–456.

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, 39(2):390–431.

Schaeck, K., Cihak, M., and Wolfe, S. (2009). Are Competitive Banking Systems More Stable? *Journal of Money, Credit and Banking*, 41(4):711–734.

Sealey, C. W. J. and Lindley, J. T. (1977). Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions. *The Journal of Finance*, 32(4):1251–1266.

Shimer, R. (2005). The Cyclical Behavior of Equilibrium and Vacancies Unemployment. *American Economic Review*, 95(1):25–49.