University College London

Doctoral Thesis

Dynamic Contracts and Labour Market Frictions

Thibaut Lamadon

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of philosophy

in the

Department of Economics

August 27, 2014

Declaration of Authorship

I, Thibaut Lamadon, declare that this thesis titled "Dynamics Contracts and Labour Frictions", and the work presented are my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: ____________________

Date: 30/07/2014

Abstract

This thesis studies the effect of repeated and long term relationships between actors engaged in economic markets. Firms hire workers for long periods and offer contracts that evolve over time, and where the history shared with the worker might affect future payments. This thesis shows that understanding the nature and implications of such relationships is central to correctly measure the realized allocation in the market and predict the effects of changes in labour policies.

The opening chapter is a theoretical contribution to the repeated games literature. It demonstrates how differences in time preferences between players can be used to sustain equilibrium payoffs that are unattainable under identical discount parameters. This reveals how rich inter-temporal strategies can be utilized to sustain improbable transfers between individuals.

The second chapter embeds such a relationship inside an equilibrium where actors randomly meet with each other. It contributes to the literature on labour markets with friction by demonstrating how widely available matched employer-employee data can be used to recover the production function in the economy as well as the assignment of workers to firms. This has important implications for the effectiveness of policies aiming at reallocating workers to more productive jobs.

In the final chapter, workers are risk averse and productivity is uncertain. I show that in this context firms choose to offer partial insurance contracts to their workers. The repeated interactions between the firm and the worker are fundamental to understanding how employers choose to transmit part of the uncertainty to the workers. I estimate the model on Swedish data and evaluate the effects of a hypothetical progressive tax aimed at reducing income inequality and uncertainty. The exercise reveals that firms will respond to the policy by transferring more risk to the employees negating around 30% of the direct effect of the policy.

Acknowledgements

I first want to thank my supervisors, Jeremy Lise, Costas Meghir and Jean-Marc Robin. When I came to UCL to study labour economics I was lucky to find this group from which I learned enormously. The department of economics at UCL and the the Institute for Fiscal Studies developed my appetite for the combination of empirically driven and theoretically sound research. Among the many faculties I have interacted with over the years I want to particularly thank Marco Bassetto, Richard Blundell, Melanie Morten, Mariacristana De Nardi, Eric French, Jan Eeckhout, Antonio Guarino, Guy Laroque, Lars Nesheim and Nicola Pavoni. I also want to thank directly the Institute for Fiscal Studies, the Centre for Microdata Methods and Practice and the The Economic and Social Research Council for funding me through my years on the PhD program. I am also thankful to the IFAU in Sweden for giving me access to the Swedish matched employer-employee records.

I thank Yale University and the Cowles Foundation for welcoming and sponsoring me for a year as a visiting scholar in 2012-13. This gave me the chance to interact with a new group of academics and allowed me to add breadth to my understanding of the field. I want to particularly thank G. Moscarini and J. Altongi for their continuous support. I thank the N. Daw lab in the Psychology department at NYU for letting me work there during my numerous trips to New York City. I also thank the NYU Economics department, and Gianluca Violante and Edouard Schall for all the comments and help in the past few years. I am extremely grateful to my examiners Philipp Kircher and Fabien Postel-Vinay for the time they invested in reading this thesis and providing me with detailed and extremely valuable comments.

I want to thank the numerous visitors at UCL and faculty members I have met on the job market this last fall for their interest and feedback in my research as well as all participants at conferences and seminars where I have had the chance to present my work: University of Minnesota, Stanford University, Northwestern University, University of Pennsylvania, University of Chicago, New York University, University of Maryland, University of Essex, the Cowles conference, Sciences-Po Paris, Sandberg conference on sorting. At both Yale and UCL, I thank all the other students I have had the chance to interact with: Aaron, Alan, Ali, Andreas, Ben, Cathy, Chris, Corrina, Dan, Dylan, Emily, Jan, Marc, Michael, Neele, Nicolas, Ran, and Vitorrio. I thank Yves Guéron and Caroline Thomas with whom I co-authered my very first paper in economics. I am extremely tankful to my great friends Florian Oswald with whom I have shared an office, countless lunches and many great evenings, and Italo Lopez, my colleague, flatmate and beer partner at our local pub. I want to give a special mention to Sara Constantino for her support over all this years, her tolerance for my workaholism, our rich discussions over a very wide set of topics and her enthusiasm and good spirit when I was in desperate need of a break!

Finally and most importantly, I want to thank my family. They have supported me and allowed me to focus on my work while providing an often needed stable environment to go back to and recharge. They have done so without ever asking for anything in return. I can't really describe how lucky I feel to have Georges and Patricia as my parents and Agathe as my sister. I feel deeply indebted to them.

Contents

Abstract **2**

Acknowledgements **3**

Introduction **7**

1 Repeated Games with One-Dimensional Payoffs and Different Discount factors **11**

1.1 Introduction 12

1.2 Lowest Equilibrium Payoffs 17

1.3 Conclusion 29

Bibliography 30

Appendix 31

1.A Two players with same discount factor 31

1.B Generalization 33

2 Identifying Sorting with On-the-job Search **35**

2.1 Introduction 36

2.2 The model 37

2.3 Properties of the equilibrium 40

2.4 Identification 45

2.5 Conclusion 52

Bibliography 53

Appendix 53

2.A More equilibrium properties 54

2.B Proof for Proposition 1 55

3 Productivity Shocks, Dynamic Contracts and Income Uncertainty 58

3.1 Introduction 59

3.2 Earnings dynamics and participation 63

3.3 The model 67

3.4 Estimation 81

3.5 Empirical implications 87

3.6 Conclusion 90

Bibliography 91

Appendix 95

3.A Data 95

3.B Proofs 96

3.C Additional Data information 108

3.D Model extensions 108

Note on co-authored work 111

Introduction

Many markets are characterized by long term relationships between actors: Firms hire workers for long periods of time and it is common for them to offer contracts with “stock options” and vesting periods, contracts indexed on performance, or even to commit to career advancements; Big retailers try to lock customers in with fidelity programs; In markets for life-insurance, monthly payments have to reflect past, current and expected future customer health condition, and companies offer very complicated dynamic prices. Understanding such pricing mechanisms, and how they affect the allocation of resources, involves analyzing both the process that brings actors together in the market, as well as the strategic interactions between these actors. In this thesis, I develop methods to address these questions both theoretically and empirically and apply these novel methods to labour markets.

The first chapter, published in *Games and Economic Behavior* with Yves Guéron and Caroline Thomas, is a theoretical contribution to the repeated games literature. The paper demonstrates how in a repeated game, differences in time preferences between players can be used to sustain equilibrium payoffs that are unattainable under identical discount parameters. Relatively patient players can make credible punishment threats to less patient players that consist of back-loading pay-offs. We show how this type of punishment strategy can be used to sustain otherwise unreachable payoffs. This allows us to relax the usual full dimension assumption necessary for the folk theorem to apply. This paper reveals how rich inter-temporal strategies can be used to sustain improbable transfers between individuals. Methods and results from the repeated game literature have important implications for the study of large markets, as shown in the following chapters.

In my second chapter, co-authored with Jeremy Lise, Costas Meghir and Jean-Marc Robin, we address the question of how workers are assigned to firms in the economy and how search frictions constrain this assignment. We develop a model with two-sided heterogeneity, production complementarities and a contract setting where workers' wages are determined by their productivity, the productivity of their employers and also their employment histories. The wage paid to a worker at a given point is a sub-game perfect equilibrium of a game played between the firm and the worker. The long term relationship between firm and worker and the repetitive nature of these interactions are central to understanding the price dispersion observed in the data, particularly among similar workers, even in similar firms. We provide a constructive identification proof of how wage data, firm size and co-worker information can be used to non- parametrically recover the production function, the assignment distribution over unobserved worker and firm types, as well as the cost to firms of creating new jobs. This paper is an important contribution to our knowledge about both the empirical content of matched employer-employee data and how well models with long-term contracts perform at matching observed wage and employment dynamics.

In the final chapter I examine the sources of earning uncertainty faced by workers in the labour market. The data tells us that a portion of earnings uncertainty is shared by co-workers at the firm level and that job losses and transitions are important sources of earning variation. In order to understand how productivity shocks are transmitted into earning and employment uncertainty, I develop an equilibrium model with search frictions, worker risk aversion and worker and firm shocks. In the model, firms optimally choose how the wage contract transmits productivity shocks to wages. I show theoretically that the presence of rents due to search frictions, together with the incentive problem due to workers' private search, result in an optimal contract that smoothly tracks underlying productivity. The repeated interactions between firm and worker are fundamental to understanding how employers choose to transmit part of the uncertainty to the workers. This represents a departure from perfectly competitive markets since worker shocks are only partially transmitted and workers are not fully shielded from firm shocks. I estimate the model with matched employer-employee data from Sweden. Using information about earnings shocks shared by co-workers, I am able to disentangle firm-specific and worker- specific shocks. Preliminary estimates suggest that firm level shocks are responsible for about 20% of permanent income fluctuations; the remaining fluctuations are accounted for by individual level shocks (30% to 40%) and job mobility (40% to 50%). Pass-through estimates reveal that the wage contract attenuates 80% of individual productivity shocks but transmits 30% of firm productivity fluctuations. Moreover, the model can be used to evaluate the effect of labour policies. In the paper, I look at the effects of a hypothetical progressive tax aimed at reducing income inequality and uncertainty. The exercise reveals that firms will respond to the policy by transferring more risk to the workers. In equilibrium, 30% of the direct effect of the policy is negated out by firms responding by posting riskier wage contracts.

The first chapter looks theoretically at the repeated interactions between actors, and contributes to our understanding about the kinds of equilibrium relationships that can be sustained over long term strategic interactions. The second chapter embeds a sub-game perfect equilibrium inside an equilibrium where actors randomly meet with each other. In the paper, we develop in detail how data can be used to recover the parameters of the model. In the final chapter, workers are risk averse and productivity is uncertain. I show that in this context firms choose to offer partial insurance contracts to their workers. I hope that the findings presented in the thesis will help future research better understand price formation in frictional labour markets, but also in other markets where long term relationships are an important feature of the data, such as insurance markets, financial markets and goods market.

Chapter 1

Repeated Games with One-Dimensional Payoffs and Different Discount factors

This chapter is based on joint work with Yves Gueron and Caroline Thomas.

1.1 Introduction

For the folk theorem to hold with more than two players, it is necessary to have the ability to threaten any single player with a low payoff, while also offering rewards to the punishing players. In assuming full dimensionality of the convex hull of the set of feasible stage-game payoffs, Fudenberg and Maskin (1986) guarantee that those individual punishments and rewards exist. Abreu, Dutta, and Smith (1994) show that the weaker NEU condition ("nonequivalent utilities"), whereby no two players have identical preferences in the stage-game, is sufficient for the folk theorem to hold.

When the NEU condition fails, players that have equivalent utilities can no longer be individually punished in equilibrium. Wen (1994) introduces the notion of *effective minmax* payoff, which takes into account the fact that when a player is being minmaxed, another player with equivalent utility might unilaterally deviate and best respond. The effective minmax payoff of a player cannot be lower than his individual minmax payoff (when NEU is satisfied, they coincide), and Wen shows that when NEU fails it is the effective minmax that constitutes the lower bound on subgame-perfect equilibrium payoffs. He establishes the following folk theorem: when players are sufficiently patient, any feasible payoff vector can be supported as a subgame-perfect equilibrium, provided it dominates the effective minmax payoff vector. We show that this can be relaxed by allowing for unequal discounting.

As pointed out by Lehrer and Pauzner (1999), when players have different discount factors, the set of feasible payoffs in a two-player repeated game is

	L	R
T	3,1	0,0
B	0,0	1,3

Figure 1.1: Battle of the sexes

typically larger and of higher dimensionality than the set of feasible stage-game payoffs.[1] In a particular three-player game in which two players have equivalent utilities, Chen (2008) illustrates how with unequal discounting payoffs below the effective minmax may indeed be achieved in equilibrium for one of the players.

In this chapter, we explore the notion that unequal discounting restores the ability to punish players individually in an n-player game where all players have equivalent utilities. Our result is stronger than Chen's as we show that all players can be hold down to their individual minmax payoff in equilibrium. Moreover we argue that our result holds for all possible violations of NEU. We find that a small difference in the discount factors suffices to hold a player to his individual minmax for a certain number of periods while still being able to reward the punishing players. For discount factors sufficiently close to one, any strictly individually rational payoff, including those dominated by the effective minmax payoff, can be obtained as the outcome of a subgame-perfect equilibrium with public correlation, restoring the validity of the folk theorem.

Although our result is stated for games where all players have equivalent utilities, we conjecture that it extends to weaker violations of NEU, as long as any two players with equivalent utilities have different discount factor. The intuition behind this conjecture is that following Abreu, Dutta, and Smith (1994) we could design specific punishments for each group of players with equivalent utilities and use the difference in discount factors within each group to enforce those specific punishments.

	L	R
T	1,1,1	0,0,0
B	0,0,0	0,0,0

C

	L	R
T	0,0,0	0,0,0
B	0,0,0	1,1,1

D

Figure 1.2: A stage game with one-dimensional payoffs

1.1.1 An Example

Consider the stage-game in Figure 1.2, where Player 1 chooses rows, Player 2 columns and Player 3 matrices. This stage-game is infinitely repeated and the players evaluate payoff streams according to the discounting criterion. When the players share a common discount factor $\delta < 1$, Fudenberg and Maskin (1986, Example 3) show that any subgame-perfect equilibrium yields a payoff of at least 1/4 (the effective minmax) to each player, whereas the individual minmax payoff of each player is zero.[2] The low dimensionality of the set of stage-game payoffs weakens the punishment that can be imposed on a player as another player with equivalent utility can deviate and best respond. The inability to achieve subgame-perfect equilibrium payoffs in $(0, 1/4)$ means that the "standard" folk theorem fails in this case.[3]

We show however that if all three players have different discount factors, there exists a subgame-perfect equilibrium in which the payoff to each player is arbitrarily close to zero, the individual minmax, provided that the discount factors are sufficiently close to one. Any payoff in the interval $(0, 1/4)$ can then be achieved in equilibrium, restoring the validity of the folk theorem in the context of this game.

[1] Mailath and Samuelson (2006, Remark 2.1.4) present a simple example to show how the set of feasible payoffs can increase when allowing for different discount factors. Consider the game of battle of the sexes depicted in Figure 1.1 and assume that players have different discount factors, $\delta_1 > \delta_2$. Consider an outcome in which (B, R) is played for T periods while (T, L) is played in subsequent periods. That is, first the less patient player is favored while the more patient player is rewarded subsequently. The playoffs to player 1 and 2 from this outcome are $(1 - \delta_1^T) + 3\delta_1^T$ and $3(1 - \delta_2^T) + \delta_2^T$, which is outside the convex hull of the set $\{(3, 1), (0, 0), (1, 3)\}$ because $\delta_1 > \delta_2$.

[2] For example, when Player 1 plays T and Player 2 plays R, Player 3 gets a payoff of 0 whether he plays C or D.

[3] One may not be too concerned about our inability to achieve low payoffs. However if the game of Figure 1.2 is part of a more general game then our inability to reach low payoffs (that is, to punish players) might reduce the scope for cooperation in the more general game.

1.1.2 Notation

We consider an n-player repeated game, where all players have equivalent utilities. We normalize payoffs to be in $\{0,1\}$ and let each player's individual minmax payoff be zero.[4] We use public correlation to convexify the payoff set, although we argue later that this assumption can be dispensed with. Players have different discount factors, and are ordered according to their patience level: $0 < \delta_1 < \cdots < \delta_{n-1} < \delta_n < 1$.[5] We use an exponential representation of discount factors: $\forall\, i,\ \delta_i := e^{-\Delta\rho_i}$, where $\Delta > 0$ could represent the length of time between two repetitions of the stage game. As $\Delta \to 0$, all discount factors tend to one. The ρ's are strictly ordered: $0 < \rho_n < \cdots < \rho_2 < \rho_1$. We assume that the stage game has a (mixed) Nash equilibrium which yields a payoff $Q < 1$ to all players.[6]

We summarize our assumptions about the game and introduce a notation for the lowest subgame-perfect equilibrium payoff of a player i in the following definitions:

Definition 1. *Let $\Gamma(\Delta)$ be the set of n-player infinitely repeated games such that:*

(i) The set of stage-game payoffs is one-dimensional and all players receive the same payoff in $\{0,1\}$.

(ii) The stage game has a mixed-strategy Nash equilibrium which yields a payoff of $Q < 1$ to all players.

(iii) Each player's pure action individual minmax payoff is zero.

(iv) Players evaluate payoff streams according to the discounting criterion, and discount factors are strictly ordered: $0 < \delta_1 < \cdots < \delta_n < 1$, where $\delta_i := e^{-\Delta\rho_i}$.

[4] We only use two payoffs as we only need to consider the minmax payoff and the maximum possible payoff.

[5] Note that the result no longer holds if several but not all players have the same discount factor. We address this point in Section 1.A.

[6] For example in the game of Figure 1.2, the mixture $\left\{ (1/2,1/2), (1/2,1/2), (1/2,1/2) \right\}$ is a Nash equilibrium that yields a payoff of 1/4.

Note that the stage game of Figure 1.2 satisfies assumptions A(i) to A(iii) of Definition 1.

Definition 2. *We denote by a_i the lowest subgame-perfect equilibrium payoff of Player i in a game $G_\Delta \in \Gamma(\Delta)$.*

For given discount factors, the existence of the $(a_i)_{i=1,\dots,n}$ is ensured by the compactness of the set of subgame-perfect equilibrium payoffs (see Fudenberg and Levine (1983, Lemma 4.2)).

1.1.3 Main Result and Outline of the Proof

Our main result, Theorem 1, states that for games in $\Gamma(\Delta)$, the lowest subgame-perfect equilibrium payoff of each player goes to zero (the common individual minmax payoff) as discount factors tend to one:

Theorem 1. *Consider an n-player infinitely repeated game $G_\Delta \in \Gamma(\Delta)$. Then $a_i \in O(\Delta)$ for all i.*[7]

Theorem 1 states that for discount factors sufficiently close to one (that is for Δ sufficiently close to zero), the lowest subgame-perfect equilibrium payoff of each player i, a_i, is arbitrarily close to zero. We do not provide a full characterization of the set of subgame-perfect equilibrium payoffs but note that any feasible and strictly individually rational payoff is a subgame-perfect equilibrium payoff. In recent work, Sugaya (2010) characterises the set of perfect and public equilibrium payoffs in games with imperfect public monitoring when players have different discount factors, under a full-dimensionality assumption.

To prove Theorem 1, we first show that when stage-game payoffs are identical, the lowest subgame-perfect equilibrium payoffs are ordered according to the discount factors (Lemma 1). A player's lowest subgame-perfect equilibrium payoff cannot be below that of another player who is less patient. We then show that the lowest subgame-perfect equilibrium payoffs of the two

[7] That is, $\exists\, M \geq 0$ and $\Delta^* > 0$ such that $a_i \leq M \cdot \Delta$ for $\Delta \leq \Delta^*$.

most patient players (Player $n-1$ and Player n) are arbitrarily close to each other when discount factors tend to one (Lemma 2). This is done by explicitly constructing a subgame-perfect equilibrium of the repeated game.

In a similar way, we then construct a set of subgame-perfect equilibria (one for each player $i \in \{2, \ldots, n-1\}$) (Lemma 3) and use those to bound the distance between the lowest subgame-perfect equilibrium payoffs of players i and $i-1$ (Lemma 4). We then show by induction that the lowest subgame-perfect equilibrium payoffs of *any* two players are arbitrarily close to each other as discount factors tend to one (Lemma 5). Finally we show that Player 1's lowest subgame-perfect equilibrium payoff can be made arbitrarily close to zero as discount factors tend to one (Lemma 6). We are then able to conclude and prove Theorem 1.

Note that the assumption of strictly different discount factors cannot be dispensed with. In particular our result does not hold when some but not all player share a common discount factor. In a similar fashion to Fudenberg and Maskin (1986, Example 3), we construct a four-player example where the stage game satisfies assumptions A(i) to A(iii) but where the two "intermediate" players share a common discount factor. That is we have $\delta_1 < \delta_2 = \delta_3 < \delta_4$. This example is presented in Section 1.A.

1.2 Lowest Equilibrium Payoffs

1.2.1 Strategy Profiles and Incentive Compatibility Constraints

To prove Theorem 1, we explicitly construct several subgame-perfect equilibria of the repeated game. To do so, we consider strategy profiles that give a constant expected stage-game payoff between zero and one (using public correlation) to all players for a given number of periods, and then stage-game payoffs of one forever:

Definition 3. *Let $\sigma(\mu, \tau, i)$ be the strategy profile such that:*

(i) For τ periods, in each stage-game, players use a public correlating device

to generate an expected payoff of μ. When the public correlating device generates a payoff of zero, players minmax Player i.

(ii) In all subsequent periods $t > \tau$, players play an action profile yielding a stage-game payoff of 1 to each player.

(iii) During the first τ periods, deviations by Player i are ignored. After that, if Player i deviates from the equilibrium path, players play a subgame-perfect equilibrium which gives Player i his lowest possible payoff, a_i.

(iv) If a deviation by Player $j \neq i$ occurs at any time, players then play a subgame-perfect equilibrium which gives Player j his lowest possible payoff, a_j.

Assuming that the correlating device generates a payoff of zero at $t = 0$, a player $j \neq i$ will not have an incentive to deviate from $\sigma(\mu, \tau, i)$ if:[8,9]

$$(1 - \delta_j) + \delta_j a_j \leq \delta_j \left((1 - \delta_j^{\tau-1})\mu + \delta_j^{\tau-1} \right), \tag{1.1}$$

which can be rewritten as

$$\delta_j^{\tau} \geq \frac{1 - \delta_j + \delta_j a_j - \delta_j \mu}{1 - \mu}. \tag{1.2}$$

To prove Theorem 1, we show that there exists a "low" μ and a large τ such that for Δ sufficiently close to zero, the strategy profile $\sigma(\mu, \tau, i)$ is subgame perfect, that is, we show that (1.2) is satisfied for any $j \neq i$. To do so, we identify the player with the tightest incentive compatibility constraint as j_i^* and find the largest τ such that (1.2) is satisfied for Player j_i^* (Lemma 3). Notice that Player j_i^* is not necessarily the player with the lowest discount factor. By a "low" μ we mean that μ must be close to a_{i-1}. To this end, we

[8] First note that zero is the lowest possible stage game payoff and so if it is enforceable all other payoffs will be. Second the strategy starts by giving zeros and ones and then rewards the players with ones forever, so the tightest incentive compatibility constraint will be when $t = 0$ as for $t > 0$ players are closer to getting ones for ever.

[9] The left-hand side of (1.1) is the payoff to Player j if he deviates: he get an instantaneous payoff of 1 followed by a repeated game payoff of a_j. If Player j follows the strategy he gets a payoff of zero today, followed by $\tau - 1$ periods during which he gets an expected payoff of μ, after which he receives a payoff of one in each period.

choose a stage-game payoff μ_i that is slightly above a_{i-1}:

Definition 4. *For all $i \in \{1, ..., n\}$, let μ_i be such that:*[10]

$$\mu_i = \begin{cases} a_{i-1} + \frac{1-\delta_1}{\delta_1} & \text{if } 2 \leq i \leq n, \\ 0 & \text{if } i = 1. \end{cases}$$

To illustrate, consider a player i with intermediate patience, such that $1 < i < n$. The strategy profile $\sigma(\mu, \tau, i)$ does not give him an opportunity to deviate, as he is being minmaxed when payoffs of zero are generated. For this reason, that strategy profile can be thought of as the other players colluding against player i. Lowering the payoff to player i from that strategy profile may conflict with making it incentive compatible both for players that are more and less patient than him. Players less patient than i must get a payoff sufficiently higher than their lowest SPE payoff, and players more patient than i must be promised payoffs of 1 soon enough to make them accept an early stream of low payoffs. We show that these constraints can be reconciled with keeping player i's payoff very close to the lowest equilibrium payoff of the player just less patient than him.

1.2.2 Proof of Theorem 1

In a first step towards Theorem 1 we now show that the lowest subgame-perfect equilibrium payoffs are ordered according to the discount factors (Lemma 1), and that Player n's lowest subgame-perfect equilibrium payoff is arbitrarily close to Player $n-1$'s for Δ close enough to zero (Lemma 2).

Lemma 1. $\forall i \in \{2, \ldots, n\}$, $a_{i-1} \leq a_i$.

Proof. The main idea is to find a stream of payoffs $(z_t)_{t=0,\ldots,\infty}$ in $[0,1]^{\mathbb{N}}$ that minimizes Player i's average discounted payoff, given Player $i-1$ is guaranteed his lowest subgame-perfect equilibrium payoff at each stage. By definition, the

[10]Note that for all i and for Δ sufficiently close to zero, $\mu_i \leq 1$. Indeed, $\mu_i \leq Q + \frac{1-\delta_1}{\delta_1} \to_{\Delta \to 0} Q < 1$.

resulting average discounted payoff for Player i cannot be greater than a_i. We show that the constraints imposed by Player $i-1$'s lowest subgame-perfect equilibrium payoff must all be binding and that $z_t = a_{i-1}$, $\forall t \geq 0$.

Formally, we solve the following minimization problem:

$$\min_{(z_t)_{t=0,\dots,\infty}\in[0,1]^{\mathbb{N}}} (1-\delta_i)\sum_{t=0}^{\infty}\delta_i^t\, z_t \tag{1.3}$$

subject to

$$(1-\delta_{i-1})\sum_{t=s}^{\infty}\delta_{i-1}^{t-s}\, z_t \geq a_{i-1}, \qquad \forall s \geq 0 \tag{1.4}$$

We show by induction that all constraints in (1.4) will be binding, which implies that $z_s = a_{i-1}$, $\forall s \geq 0$. Our induction hypothesis is that the constraints in (1.4) must bind for $s = 0, \dots, \tau$ and therefore, that the minimization problem (1.3) subject to the constraints (1.4) can be rewritten as:

$$\min_{(z_t)_{t=\tau,\dots,\infty}\in[0,1]^{\mathbb{N}}} \lambda_{\tau-1}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\left(\sum_{t=\tau+1}^{\infty}\delta_i^{\tau}\left(\delta_i^{t-\tau}-\delta_{i-1}^{t-\tau}\right)\, z_t\right) \tag{1.5}$$

subject to

$$(1-\delta_{i-1})\sum_{t=s}^{\infty}\delta_{i-1}^{t-s}\, z_t \geq a_{i-1}, \qquad \forall s \geq \tau+1 \tag{1.6}$$

where the function λ_τ is recursively defined by

$$\lambda_0\left(a_{i-1},\delta_{i-1},\delta_i\right) = (1-\delta_i)\,\frac{a_{i-1}}{1-\delta_{i-1}}$$

and

$$\lambda_\tau\left(a_{i-1},\delta_{i-1},\delta_i\right) = \lambda_{\tau-1}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\,\delta_i^{\tau} + (\delta_i-\delta_{i-1})\,\frac{a_{i-1}}{1-\delta_{i-1}}.$$

Initialization: $\tau = 0$ The first constraint is the only constraint featuring z_0 and can be rewritten as $z_0 \geq \frac{a_{i-1}}{1-\delta_{i-1}} - \sum_{t=1}^{\infty}\delta_{i-1}^t\, z_t$. Moreover, z_0 enters with a positive coefficient in the objective function, therefore, the first constraint must be binding. The constraint is then used to eliminate z_0 from the objective function: the minimization problem (1.3) subject to (1.4) can therefore be written in the following way:

$$\min_{(z_t)_{t=1,\ldots,\infty}\in[0,1]^{\mathbb{N}}} (1-\delta_i)\left(\frac{a_{i-1}}{1-\delta_{i-1}} + \sum_{t=1}^{\infty}\left(\delta_i^t - \delta_{i-1}^t\right) z_t\right)$$

subject to

$$(1-\delta_{i-1})\sum_{t=s}^{\infty}\delta_{i-1}^{t-s} z_t \geq a_{i-1}, \qquad \forall s \geq 1$$

This verifies (1.5) and (1.6).

Induction We assume that our minimization problem can be rewritten as (1.5) subject to (1.6) for some $\tau > 1$. Because $\delta_i > \delta_{i-1}$, $z_{\tau+1}$ enters with a positive coefficient in the objective function and $z_{\tau+1}$ only appears in the constraint $z_{\tau+1} \geq \frac{a_{i-1}}{1-\delta_{i-1}} - \sum_{t=\tau+2}^{\infty}\delta_{i-1}^{t-(\tau+1)} z_t$, this constraint will be binding and the objective function can be rewritten by substituting for $z_{\tau+1}$ as follows:

$$\begin{aligned}
&\lambda_{\tau-1}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\left(\sum_{t=\tau+1}^{\infty}\delta_i^{\tau}\left(\delta_i^{t-\tau} - \delta_{i-1}^{t-\tau}\right) z_t\right) \\
&= \lambda_{\tau-1}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\left(\delta_i^{\tau}\left(\delta_i - \delta_{i-1}\right)\left(\frac{a_{i-1}}{1-\delta_{i-1}} - \sum_{t=\tau+2}^{\infty}\delta_{i-1}^{t-(\tau+1)} z_t\right)\right) \\
&\quad + (1-\delta_i)\sum_{t=\tau+2}^{\infty}\delta_i^{\tau}\left(\delta_i^{t-\tau} - \delta_{i-1}^{t-\tau}\right) z_t \\
&= \lambda_{\tau}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\sum_{t=\tau+2}^{\infty}\left(\delta_i^{\tau}\left(\delta_i^{t-\tau} - \delta_{i-1}^{t-\tau}\right) - \delta_i^{\tau}\left(\delta_i - \delta_{i-1}\right)\delta_{i-1}^{t-(\tau+1)}\right) z_t \\
&= \lambda_{\tau}\left(a_{i-1},\delta_{i-1},\delta_i\right) + (1-\delta_i)\left(\sum_{t=\tau+2}^{\infty}\delta_i^{\tau+1}\left(\delta_i^{t-(\tau+1)} - \delta_{i-1}^{t-(\tau+1)}\right) z_t\right),
\end{aligned}$$

where the first equality is obtained by substituting for $z_{\tau+1}$ and the other equalities are obtained by grouping the terms in z_t ($t \geq \tau+2$) together. Thus (1.5) and (1.6) hold for $\tau+1$ also.

This concludes the proof by induction and so all constraints in (1.4) must bind: $(1-\delta_{i-1})\sum_{t=s}^{\infty}\delta_{i-1}^{t-s} z_t = a_{i-1}$, $\forall s \geq 0$. We now show that this implies that $z_s = a_{i-1}$, $\forall s \geq 0$. Consider the constraint for some $s \geq 0$:

$$\begin{aligned} a_{i-1} &= (1-\delta_{i-1}) \sum_{t=s}^{\infty} \delta_{i-1}^{t-s}\, z_t \\ &= (1-\delta_{i-1}) \left\{ z_s + \delta_{i-1} \sum_{t=s+1}^{\infty} \delta_{i-1}^{t-(s+1)}\, z_t \right\} \\ &= (1-\delta_{i-1}) \left\{ z_s + \frac{\delta_{i-1}}{1-\delta_{i-1}}\, a_{i-1} \right\}, \end{aligned}$$

where the last inequality holds because the constraint is binding for $s+1$. This implies that $z_s = a_{i-1}$, $\forall s \geq 0$.

Given the constraints imposed on stage-game payoffs by player $i-1$'s lower subgame-perfect equilibrium bound, the lowest average discounted payoff which can be given to player i is a_{i-1}. We therefore have $a_{i-1} \leq a_i$. □

Lemma 2. $|a_n - a_{n-1}| \in O(\Delta)$.

Proof. Consider the strategy profile $\sigma(\mu_n, \infty, n)$, where $\mu_n = a_{n-1} + \frac{1-\delta_1}{\delta_1}$. We are going to show that this constitutes a subgame-perfect equilibrium.

First, note that in a period in which the public correlating device generates a payoff of one, no player has a one-shot profitable deviation. Secondly, because Player n is being minmaxed in a period in which the public correlating device generates a payoff of zero, he doesn't have a profitable one-shot deviation. Thirdly, because punishment phases consist of subgame-perfect equilibrium strategies, no player has a profitable one-shot deviation during one of those. Thus, to verify that $\sigma(\mu_n, \infty, n)$ is subgame perfect, we only need to check that players $i \leq n-1$ do not have profitable one-shot deviations when the public correlating device generates a payoff of zero.

A deviation from Player $i \leq n-1$ leads at most to a one-off gain of one followed by a payoff of a_i forever. Therefore, there is no one-shot profitable deviation if $(1-\delta_i) + \delta_i a_i \leq \delta_i \left(a_{n-1} + \frac{1-\delta_1}{\delta_1}\right)$, where the right-hand-side is the repeated game payoff to Player i if the public correlation device indicates a zero payoff action profile in that period. This inequality is always satisfied for $i \leq n-1$ as $a_i \leq a_{n-1}$ (Lemma 1) and as $\frac{1-\delta_i}{\delta_i} \leq \frac{1-\delta_1}{\delta_1}$.

By definition of a_n, and by Lemma 1, we have that $a_{n-1} \leq a_n \leq a_{n-1} + \frac{1-\delta_1}{\delta_1}$. We conclude the proof by noting that $a_n - a_{n-1} \leq \frac{1-\delta_1}{\delta_1}$ and that $\frac{1-\delta_1}{\delta_1} \in O(\Delta)$. □

We have shown that the lowest subgame-perfect equilibrium payoffs of the two most patient players are arbitrarily close as Δ tends to zero. The intuition behind this result is that all players can collude against Player n by minmaxing him whenever the public correlating device generates a payoff of zero. Since Player $n-1$ is the most patient of the colluding players and since lowest subgame-perfect equilibrium payoffs are ordered according to discount factors, his lowest subgame-perfect equilibrium will determine by how much Player n's equilibrium payoff can be pushed down.

We now show that the lowest subgame-perfect equilibrium payoffs of *any* two players are arbitrarily close to each other as Δ tends to zero (Lemma 5). We start by identifying bounds on Player $i > 1$'s lowest subgame-perfect equilibrium payoff. To do this, we find the largest time $\tau \geq 1$ such that the strategy profile $\sigma(\mu_i, \tau, i)$ is a subgame-perfect equilibrium and compute its equilibrium payoff for Player i. We then prove Lemma 5 by induction.

First, we introduce some useful notation. For every player $i \in \{1, \ldots, n-1\}$, define

$$N_+^i := \{j > i : 1 - \delta_j + \delta_j a_j - \delta_j \mu_i > 0\}.$$

When proving that for a particular τ, $\sigma(\mu_i, \tau, i)$ is a subgame-perfect equilibrium, N_+^i should be thought of as the set of players for whom profitable deviations might exist depending on the value of τ. That is, N_+^i is the set of players for whom the right-hand side of (1.2) (when replacing μ with μ_i) is strictly positive. We will therefore choose τ to satisfy the no-deviation constraints of all players in N_+^i. When N_+^i is not empty, we identify the player from this set with the tightest constraint as j_i^* and we define $\widetilde{t}_i$ as follows:

$$j_i^* := \arg\min_{j \in N_+^i} \frac{\log\Big((1-\delta_j+\delta_j a_j - \delta_j \mu_i) \,/\, (1-\mu_i) \Big)}{\log \delta_j},$$

$$\widetilde{t}_i := \frac{\log\Big(\Big(1-\delta_{j_i^*}+\delta_{j_i^*} a_{j_i^*} - \delta_{j_i^*} \mu_i\Big) \,/\, (1-\mu_i) \Big)}{\log \delta_{j_i^*}}.$$

Let $t_i^* := \left\lfloor \widetilde{t}_i \right\rfloor$ be the largest integer smaller or equal than $\widetilde{t}_i$ and define $r_i \in (0,1)$ to be the fractional part of $\widetilde{t}_i$:

$$r_i := \widetilde{t}_i - t_i^*.$$

Note that t_i^* is the longest time τ such that j_i^* does not have a profitable one-shot deviation in $\sigma(\mu_i, \tau, i)$.

In Lemma 3 we show that for Δ sufficiently close to zero t_i^* is well defined and arbitrarily large and that the strategy profile $\sigma(\mu_i, t_i^*, i)$ is indeed subgame perfect.

Lemma 3. *Let $i \in \{2, ..., n-1\}$, and assume that $N_+^i \neq \emptyset$. Given j_i^*, t_i^* and μ_i, $\exists\, \Delta_i^* > 0$ such that for $\Delta \in (0, \Delta_i^*)$, $\sigma(\mu_i, t_i^*, i)$ constitutes a subgame-perfect equilibrium.*

Proof. For notational convenience, we omit the i subscript on j_i^*, $\widetilde{t}_i$, t_i^*, and r_i. First, recall that for Δ sufficiently close to zero, $\mu_i \leq 1$.[11] We now check that t^* is well defined. Note that $\exists\, \Delta_{ij} > 0$ and $\eta_{ij} < 1$ such that for $\Delta \leq \Delta_{ij}$, $\frac{1-\delta_j+\delta_j a_j-\delta_j\mu_i}{1-\mu_i} < \eta_{ij}$.[12] Because η_{ij} does not depend on Δ, this shows that $\lim_{\Delta\to 0} \widetilde{t} = \infty$ and ensures that $\exists\, \Delta_i^* > 0$ such that t^* is well defined and strictly positive for $\Delta \in (0, \Delta_i^*)$.

Because i is being minmaxed if the public correlating device generates a payoff of zero, i does not have a profitable one-shot deviation. Also, no

[11] See footnote 10.

[12] Since $a_j \leq Q$, $\frac{1-\delta_j+\delta_j a_j-\delta_j\mu_i}{1-\mu_i} \leq \delta_j \frac{Q-\mu_i}{1-\mu_i} + \frac{1-\delta_j}{1-Q-(1-\delta_1)/\delta_1}$. For any x in $[0,1)$, $\frac{Q-x}{1-x} \leq Q$, thus the right-hand-side of the previous inequality is bounded from above by $\delta_j Q + \frac{1-\delta_j}{1-Q-(1-\delta_1)/\delta_1}$, which tends to $Q < 1$ as Δ tends to zero.

player will have a profitable one-shot deviation during the punishment phases of $\sigma(\mu_i, t_i^*, i)$, as those are subgame perfect.

We now check that no player $j \neq i$ has a profitable one-shot deviation, that is, we check that (1.1) (when replacing μ with μ_i and τ with t^*) holds for all players $j \neq i$:

$$(1-\delta_j) + \delta_j a_j \leq \delta_j \left((1-\delta_j^{t^*-1})\mu_i + \delta_j^{t^*-1} \right). \tag{1.7}$$

We first check that (1.7) holds for players $j \leq i-1$ and then for players $j > i$:

(i) No deviation from player $j \leq i-1$: Note that because $\mu_i \in [0,1]$, we have that $\mu_i \leq \left(1-\delta_j^{t^*-1}\right)\mu_i + \delta_j^{t^*-1}$. In order to show that (1.7) holds, we can therefore show that $(1-\delta_j) + \delta_j a_j \leq \delta_j \mu_i$, which is equivalent to $\frac{1-\delta_j}{\delta_j} + a_j \leq a_{i-1} + \frac{1-\delta_1}{\delta_1}$. This inequality holds $\forall j \leq i-1$, as $\frac{1-\delta_j}{\delta_j} \leq \frac{1-\delta_1}{\delta_1}$ and $a_j \leq a_{i-1}$.

(ii) No deviation from player $j > i$: We can rearrange (1.7) to get

$$\delta_j^{t^*} \geq \frac{1-\delta_j + \delta_j a_j - \delta_j \mu_i}{1-\mu_i}. \tag{1.8}$$

First, note that if $j \notin N_+^i$ then j has no incentive to deviate as $\delta_j^{t^*} > 0 \geq \frac{1-\delta_j+\delta_j a_j - \delta_j \mu_i}{1-\mu_i}$. Now let $j \in N_+^i$. Since t^* has been chosen such that (1.8) is satisfied for player j^*, (1.8) is also satisfied for all other players in N_+^i, and no player $j \in N_+^i$ will have an incentive to deviate.

We conclude that for Δ sufficiently close to zero, $\sigma(\mu_i, t_i^*, i)$ is a subgame-perfect equilibrium. $\square$

Remark 1 (Dispensability of public correlation)**.** *In Lemma 3, we show that $\sigma(\mu_i, t_i^*, i)$ is a subgame-perfect equilibrium and that t_i^* goes to infinity as Δ approaches zero. Instead of using the strategy $\sigma(\mu_i, t_i^*, i)$, which relies on public correlation, we can consider a deterministic strategy that alternates between $t_{i,1}^*$ zeros and $t_{i,2}^*$ ones, where $t_{i,1}^* + t_{i,2}^* = t_i^*$ and $t_{i,2}^*/t_i^*$ is arbitrarily close to μ_i, starting with a payoff of zero. This is possible because t_i^* goes to infinity.*

Intuitively, as Δ goes to zero, such a strategy will yield a payoff to any player arbitrarily close to the payoff from $\sigma(\mu_i, t_i^*, i)$, while having a period-zero incentive compatibility constraint less stringent than (1.7) since μ_i is promised on average over the first t_i^* periods and the first period payoff is a zero. This should ensure that Lemmas 3 and 4 still hold under such a deterministic strategy. $\diamondsuit$

We now compute the payoff of player i from $\sigma(\mu_i, t_i^*, i)$ in order to bound the distance between a_i and a_{i-1}.

Lemma 4. *$\forall i \in \{2, \dots, n-1\}$, we have that either:*

(i) $\forall\, j > i$, $|a_j - a_{i-1}| \in O(\Delta)$, or
(ii) $|a_i - a_{i-1}| \in O(\Delta) + O\left(a_{j_i^*} - a_i\right)$, where $j_i^* > i$.

Proof. Again, for notational convenience, we omit the i subscript on j_i^*, t_i^* and r_i. If N_+^i is empty we directly have an indication of the distance between a_j and a_{i-1} by noting that no player $j > i$ has an incentive to deviate from $\sigma(\mu_i, \tau, i)$, irrespective of τ: if $N_+^i = \emptyset$, then $\forall\, j > i$, $0 \leq a_j - a_{i-1} \leq \frac{1-\delta_1}{\delta_1} - \frac{1-\delta_j}{\delta_j}$, which implies that $|a_j - a_{i-1}| \in O(\Delta)$.

Assume now that $N_+^i \neq \emptyset$, so that $\sigma(\mu_i, t^*, i)$ is a subgame-perfect equilibrium. We now compute Player i's payoff from $\sigma(\mu_i, t^*, i)$ and compare it with his lowest subgame-perfect equilibrium payoff. The payoff to Player i from the strategy profile $\sigma(\mu_i, t^*, i)$ is:

$$\begin{aligned}
(1-\delta_i^{t^*})\mu_i + \delta_i^{t^*} &= \mu_i + \delta_i^{t^*}(1-\mu_i) \\
&= \mu_i + \delta_i^{-r}\left(\frac{1-\delta_{j^*}+\delta_{j^*}a_{j^*}-\delta_{j^*}\mu_i}{1-\mu_i}\right)^{\frac{\rho_i}{\rho_{j^*}}}(1-\mu_i) \\
&\geq a_i,
\end{aligned}$$

where the last inequality holds because a_i is i's lowest subgame-perfect equilibrium payoff. This inequality can be rewritten as

$$\frac{a_i-\mu_i}{1-\mu_i} \leq \delta_i^{-r}\left(\frac{1-\delta_{j^*}+\delta_{j^*}a_{j^*}-\delta_{j^*}\mu_i}{1-\mu_i}\right)^{\frac{\rho_i}{\rho_{j^*}}-1}\left(\frac{1-\delta_{j^*}+\delta_{j^*}a_{j^*}-\delta_{j^*}\mu_i}{1-\mu_i}\right),$$

where $\frac{\rho_i}{\rho_{j^*}} - 1 > 0$, as $i < j^*$. Recall from the proof of Lemma 3 that for $\Delta \leq \Delta_{ij^*}$, $(1 - \delta_{j^*} + \delta_{j^*} a_{j^*} - \delta_{j^*} \mu_i) / (1 - \mu_i) < \eta_{ij^*}$, where $\eta_{ij^*} < 1$ does not depend on Δ. For $\Delta \leq \Delta_{ij^*}$, we therefore have:

$$\frac{a_i - \mu_i}{1 - \mu_i} \leq \delta_i^{-r} \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} \left(\frac{1 - \delta_{j^*} + \delta_{j^*} a_{j^*} - \delta_{j^*} \mu_i}{1 - \mu_i} \right).$$

The previous inequality can be rewritten as:[13]

$$\begin{aligned} a_i - a_{i-1} \leq \frac{1 - \delta_1}{\delta_1} + \delta_i^{-r} \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} \delta_{j^*} (a_i - a_{i-1}) + \\ \delta_i^{-r} \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} \left(1 - \delta_{j^*} + \delta_{j^*} (a_{j^*} - a_i) - \delta_{j^*} \frac{1 - \delta_1}{\delta_1} \right). \end{aligned} \tag{1.9}$$

Because

$$\lim_{\Delta \to 0} \delta_i^{-r} \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} \delta_{j^*} = \lim_{\Delta \to 0} \delta_i^{-r} \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} = \eta_{ij^*}^{\frac{\rho_i}{\rho_{j^*}} - 1} < 1,$$

there exists a $\widetilde{\Delta_i} \geq 0$ and an $R < 1$ such that for $\Delta \leq \widetilde{\Delta_i}$ we have:

$$a_i - a_{i-1} \leq \frac{1 - \delta_1}{\delta_1} + R (a_i - a_{i-1}) + R \left(1 - \delta_{j^*} + \delta_{j^*} (a_{j^*} - a_i) - \delta_{j^*} \frac{1 - \delta_1}{\delta_1} \right).$$

To conclude, note that $\frac{1-\delta_1}{(1-R)\delta_1} + \frac{R}{1-R} \left(1 - \delta_{j^*} - \delta_{j^*} \frac{1-\delta_1}{\delta_1} \right)$ is of order Δ, and that $\frac{R}{1-R} \delta_{j^*} (a_{j^*} - a_i) \in O(a_{j^*} - a_i)$, as $R < 1$ is a fixed constant. □

Recall that the difference between the two most patient players' lowest subgame-perfect equilibrium payoffs, a_n and a_{n-1}, is of order Δ (Lemma 2). Moreover in Lemma 4 we established a bound for the distance between a_{i-1} and the lowest subgame-perfect equilibrium payoff of a more patient player. We can now establish by induction that the lowest subgame-perfect equilibrium payoffs of *any* two players are arbitrarily close to each other as Δ tends to zero.

Lemma 5. *$|a_i - a_j| \in O(\Delta)$, $\forall (i, j)$.*

[13] By canceling the $1 - \mu_i$ and adding and subtracting $\delta_{j^*} a_i$ inside the term in parentheses.

Proof. By Lemma 2, we know that this result is true for $i,\, j \in \{n-1, n\}$. We now prove this result by induction. Assume that $\forall i,\, j \geq k,\, |a_i - a_j| \in O(\Delta)$. Our aim is to show that $\forall i \geq k,\, |a_i - a_{k-1}| \in O(\Delta)$.

If the first statement of Lemma 4 holds, then we have that $\forall j > k,\, |a_j - a_{k-1}| \in O(\Delta)$. Moreover, $|a_k - a_{k-1}| \leq |a_k - a_j| + |a_j - a_{k-1}|$ for any $j > k$. By induction, $|a_k - a_j| \in O(\Delta)$, thus we have $|a_k - a_{k-1}| \in O(\Delta)$.

If the second statement of Lemma 4 holds then $\exists\, k^* > k$ such that $|a_k - a_{k-1}| \in O(\Delta) + O(a_{k^*} - a_k)$. From our induction hypothesis, $|a_{k^*} - a_k| \in O(\Delta)$, which implies that $|a_k - a_{k-1}| \in O(\Delta)$. Using the triangle inequality, $\forall i \geq k,\, |a_i - a_{k-1}| \leq |a_i - a_k| + |a_k - a_{k-1}| \in O(\Delta)$.

This shows that $\forall i,\, j \geq k-1,\, |a_i - a_j| \in O(\Delta)$. □

Finally, we show that the lowest subgame-perfect equilibrium payoff of Player 1 is arbitrarily close to zero as Δ tends to zero. This is done by using a proof similar to the one of Lemma 4, and considering the strategy profile $\sigma(0, t_1^*, 1)$.

Lemma 6. $a_1 \in O(\Delta)$.

Proof. We follow the same line of reasoning as in the proof of Lemma 3 and Lemma 4, using the strategy $\sigma(0, t_1^*, 1)$. As in Lemma 3, $\sigma(0, t_1^*, 1)$ is well defined and constitutes a subgame-perfect equilibrium. Again, for notational convenience, we omit the subscript 1.

The strategy profile $\sigma(0, t^*, 1)$ yields a payoff of $\delta_1^{t^*} = \delta_1^{-r}\left(1 - \delta_{j^*} + \delta_{j^*} a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}}$ to Player 1. Because a_1 is player 1's lowest subgame-perfect equilibrium payoff, we have

$$
\begin{aligned}
a_1 &\leq \delta_1^{-r}\left(1 - \delta_{j^*} + \delta_{j^*} a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}} \\
&= \delta_1^{-r}\left(1 - \delta_{j^*} + \delta_{j^*} a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}-1}\left(1 - \delta_{j^*} + \delta_{j^*}\left(a_{j^*} - a_1\right)\right) \\
&\quad + \delta_1^{-r}\delta_{j^*}\left(1 - \delta_{j^*} + \delta_{j^*} a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}-1} a_1.
\end{aligned}
$$

Because

$$\lim_{\Delta\to 0}\delta_1^{-r}\delta_{j^*}\left(1-\delta_{j^*}+\delta_{j^*}a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}-1}=\lim_{\Delta\to 0}\delta_1^{-r}\left(1-\delta_{j^*}+\delta_{j^*}a_{j^*}\right)^{\frac{\rho_1}{\rho_{j^*}}-1}\leq\eta_{1j^*}^{\frac{\rho_1}{\rho_{j^*}}-1},$$

and $\eta_{1j^*}^{\frac{\rho_1}{\rho_{j^*}}-1}<1$ there exists an $R<1$ and $\Delta_1^*\geq 0$ such that for $\Delta\leq\Delta_1^*$ we have

$$a_1\leq R\Big(1-\delta_{j^*}+\delta_{j^*}\left(a_{j^*}-a_1\right)\Big)+Ra_1,$$

or

$$a_1\leq\frac{R}{1-R}\Big(1-\delta_{j^*}+\delta_{j^*}\left(a_{j^*}-a_1\right)\Big).$$

We know from Lemma 5 that $a_{j^*}-a_1\in O\left(\Delta\right)$, which concludes the proof, as $R<1$ does not depend on Δ. □

We are now able to prove Theorem 1:

Proof of Theorem 1. From Lemma 5 and 6, we have that $\forall i\in\{1,\ldots,n\}$, $|a_i-a_1|\in O\left(\Delta\right)$ and $a_1\in O\left(\Delta\right)$. Using the triangle inequality, $|a_i|\leq|a_i-a_1|+|a_1|\in O(\Delta)$. □

1.3 Conclusion

In this chapter, we considered the set of games where the classical folk theorem does not apply because of the low dimensionality of the set of stage-game payoffs. In such setups, it is not possible to create player-specific punishments which are necessary to sustain low values of equilibrium payoffs.

We extend the setting by allowing players to have different discount factors and prove that player-specific punishments as close as desired to the player's individual minmax can be constructed. Those punishments can be used to enforce any stage-game payoff as an equilibrium payoff. This generalizes the folk theorem to games which violate NEU but where players have different discount factors. They can also be used to yield equilibrium payoffs strictly outside the convex hull of the stage-game payoffs. However, the characterization of this multidimensional boundary for the complete equilibrium pay off set is left for future research.

In the next sections, we first show that our result does require all players to have different discount factors and does not hold if two "intermediate" players share the same discount factor. We then briefly discuss subsequent research that generalizes our result.

Bibliography

Abreu, D., P. K. Dutta, and L. Smith (1994): "The Folk Theorem for Repeated Games: A Neu Condition," *Econometrica*, 62(4), 939–948.

Chen, B. (2008): "On effective minimax payoffs and unequal discounting," *Econ. Lett.*, 100(1), 105–107.

Chen, B., and S. Takahashi (2012): "A folk theorem for repeated games with unequal discounting," *Games and Economic Behavior*, 76(2), 571 – 581.

Fudenberg, D., and D. Levine (1983): "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," *J. Econ. Theory*, 31(2), 251–268.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.

Lehrer, E., and A. Pauzner (1999): "Repeated games with differential time preferences," *Econometrica*, 67(2), 393–412.

Mailath, G. J., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press.

Sugaya, T. (2010): "Characterizing the Limit Set of PPE Payoffs with Unequal Discounting," Mimeo.

Wen, Q. (1994): "The "folk theorem" for repeated games with complete information," *Econometrica*, 62(4), 949–954.

1.A Two players with same discount factor

In this section we confirm that our result does indeed require all players to have different discount factors by means of a counter-example similar to the one presented in Section 1.1.1, but with four players. We present a particular four-player game in which player 1 and player 2 have the same discount factor $\tilde{\delta} \in (\delta_3, \delta_4)$, but such that in every stage game at least one of them is guaranteed a payoff of 1/2.

In the game of Figure 1.3, player 1 chooses a row, player 2 chooses a column, player 3 chooses between the two left matrices or the two right ones and player 4 chooses between the two top matrices or the two bottom ones. Notice that in this game, the min-max payoff of each player is 0, and there is a mixed-strategy Nash equilibrium $(1/2, 1/2, 1/2, 1)$ which yields a payoff of 1/2, so that assumptions A1 to A3 are satisfied.

0,0,0,0	1,1,1,1		1,1,1,1	0,0,0,0
0,0,0,0	1,1,1,1		1,1,1,1	0,0,0,0
0,0,0,0	1,1,1,1		0,0,0,0	0,0,0,0
0,0,0,0	1,1,1,1		1,1,1,1	1,1,1,1

Figure 1.3: A four-player stage game with one-dimensional payoffs

Let α_i denote the probability with which player i plays his first action (either top or left). The expected payoff to all players from strategy profile $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in [0,1]^4$ is

$$(1-\alpha_2)\alpha_3 + \alpha_2(1-\alpha_3)\alpha_4 + (1-\alpha_1)(1-\alpha_3)(1-\alpha_4).$$

We now show that for any stage-game action profile, at least one of player 1 or player 2 has a deviation guaranteeing him a payoff of 1/2.

Consider the payoff from a deviation for player 1. If player 1 plays the top row ($\alpha_1 = 1$), his payoff is

$$u_1^1 = (1 - \alpha_2)\alpha_3 + \alpha_2(1 - \alpha_3)\alpha_4,$$

while his payoff from playing the bottom row ($\alpha_1 = 0$) is

$$u_1^0 = (1 - \alpha_2)\alpha_3 + \alpha_2(1 - \alpha_3)\alpha_4 + (1 - \alpha_3)(1 - \alpha_4).$$

For player 1, playing the bottom row is the best deviation (bottom indeed weakly dominates top).

For player 2, the payoff from playing the left column ($\alpha_2 = 1$) is

$$u_2^1 = (1 - \alpha_3)\alpha_4 + (1 - \alpha_1)(1 - \alpha_3)(1 - \alpha_4),$$

while his payoff from playing the right column ($\alpha_2 = 0$) is

$$u_2^0 = \alpha_3 + (1 - \alpha_1)(1 - \alpha_3)(1 - \alpha_4).$$

We now show that $\max\{u_1^0, u_2^0, u_2^1\} \geq 1/2$ for any quadruple $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. First let $\beta_i = 1 - \alpha_i$. We can then rewrite u_1^0, u_2^0 and u_2^1 as $\beta_2(1 - \beta_3) + (1 - \beta_2)\beta_3(1 - \beta_4) + \beta_3\beta_4 = (1 - 2\beta_3 + \beta_3\beta_4)\beta_2 + \beta_3$, $\beta_3(1 - \beta_4) + \beta_1\beta_3\beta_4$ and $1 - \beta_3 + \beta_1\beta_3\beta_4$, respectively. As β_2 only appears in u_1^0, we can first minimize $\max\{u_1^0, u_2^0, u_2^1\}$ with respect to β_2. Moreover, u_1^0 is linear in β_2, so that it's minimum is $\beta_3 + \min(0, 1 - 2\beta_3 + \beta_3\beta_4)$.

We now notice that β_4 only appears in the expression $\beta_3\beta_4$ and that β_1 only appears in the expression $\beta_1\beta_3\beta_4$. Let $\gamma_4 = \beta_3\beta_4$ and $\gamma_1 = \beta_1\gamma_4$, our problem is equivalent to showing that the minimum for γ_1, β_3 and γ_4 such that $0 \leq \gamma_1 \leq \gamma_4 \leq \beta_3 \leq 1$ of the maximum between $\beta_3 + \min(0, 1 - 2\beta_3 + \gamma_4)$, $\beta_3 - \gamma_4 + \gamma_1$ and $1 - \beta_3 + \gamma_1$ is greater than one half.

Consider first the case when $1 - 2\beta_3 + \gamma_4 \geq 0$. Our problem is to show that $\max\{\beta_3, \beta_3 - \gamma_4 + \gamma_1, 1 - \beta_3 + \gamma_1\} \geq 1/2$ whenever $1 - 2\beta_3 + \gamma_4 \geq 0$ and

$0 \leq \gamma_1 \leq \gamma_4 \leq \beta_3 \leq 1$. Given that $\gamma_1 \leq \gamma_4$ then $\beta_3 \geq \beta_3 - \gamma_4 + \gamma_1$. First, if β_3 is the maximum of those three terms then $\beta_3 \geq 1 - \beta_3 + \gamma_1$, so that $2\beta_3 \geq 1 + \gamma_1 \geq 1$, or $\beta_3 \geq 1/2$. Second, if $1 - \beta_3 + \gamma_1$ is the maximum of those three terms then $1 - \beta_3 + \gamma_1 \geq \beta_3$, so that $\beta_3 \leq (1 + \gamma_1)/2$ and therefore $1 - \beta_3 + \gamma_1 \geq 1 - (1 + \gamma_1)/2 + \gamma_1 = (1 + \gamma_1)/2 \geq 1/2$.

Consider now the case when $1 - 2\beta_3 + \gamma_4 \leq 0$. Our problem is to show that $\max\{1 - \beta_3 + \gamma_4, \beta_3 - \gamma_4 + \gamma_1, 1 - \beta_3 + \gamma_1\} \geq 1/2$ whenever $1 - 2\beta_3 + \gamma_4 \leq 0$ and $0 \leq \gamma_1 \leq \gamma_4 \leq \beta_3 \leq 1$. Given that $\gamma_1 \leq \gamma_4$ then $1 - \beta_3 + \gamma_4 \geq 1 - \beta_3 + \gamma_1$. First if $1 - \beta_3 + \gamma_4 \geq \beta_3 - \gamma_4 + \gamma_1$ then $\beta_3 \leq 1/2 + \gamma_4 - \gamma_1/2$, so that $1 - \beta_3 + \gamma_4 \geq (1 + \gamma_1)/2 \geq 1/2$. Second if $1 - \beta_3 + \gamma_4 \leq \beta_3 - \gamma_4 + \gamma_1$ then $\beta_3 \geq 1/2 + \gamma_4 - \gamma_1/2$, so that $\gamma_1 + \beta_3 - \gamma_4 \geq (1 + \gamma_1)/2 \geq 1/2$.

Hence there is always one player amongst player 1 and player 2 who can achieve a payoff of 1/2 in the stage game. Therefore for any stage-game profile $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in [0, 1]^4$, both player 1 and player 2 are guaranteed a repeated-game payoff of at least

$$(1 - \tilde{\delta})\frac{1}{2} + \tilde{\delta}u^*,$$

where u^* is the minimum payoff attainable in any subgame-perfect equilibrium for players 1 and 2.[14] If $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ is part of an equilibrium that gives players 1 and 2 their lowest subgame-perfect equilibrium payoff we then have:

$$u^* \geq (1 - \tilde{\delta})\frac{1}{2} + \tilde{\delta}u^*,$$

so that

$$u^* \geq \frac{1}{2}.$$

1.B Generalization

In a more recent paper, Chen and Takahashi (2012) generalize our result. They aggregate the stage-game dimensionality assumption with the different

[14] Note that because they have the same stage game payoffs and the same discount factor, players 1 and 2 must have the same lowest subgame-perfect equilibrium payoff.

discount factor assumption in a dynamic non-equivalent utility assumption (DNEU). DNEU simply states that when players have equivalent utilities they must have different discount factors.[15] Chen and Takahashi (2012) dispense with the pure minmax assumption that we make and provide a more explicit construction of the dynamic player specific punishments, whereas we rely on the compactness of the equilibrium payoff set and use this to provide bounds on the difference between the lowest equilibrium payoffs of any two players. We note however that Chen and Takahashi (2012) rely on the compactness of the set of feasible repeated-game payoffs and use the lowest feasible payoff for each player without explicitly constructing them.

[15] In this chapter we considered a case in which all players have equivalent utilities, which is the most problematic case for the folk theorem. DNEU therefore reduces to having all players have different discount factors in that case.

Chapter 2

Identifying Sorting with On-the-job Search

2.1 Introduction

How does the labor market allocate workers to firms in equilibrium? Becker (1974) tells us that in a frictionless environment with complementaries in production, better workers should be assigned to better firms and that matching should be one-to-one. The data however tells us a different story where similar workers are paid differently and where firms hire many different types of workers (Mortensen, 2005). The presence of rents generated by search frictions can explain the departure from the Beckerian world, indeed if it takes time to find a new matching opportunity, workers and firms might settle for less than optimal partners, generating mismatch in equilibrium. Shimer and Smith (2003); Atakan (2006); Eeckhout and Kircher (2010) extend the result of Becker (1974) to environment with search frictions and show that the principle that stronger complementarities pushes better workers into better firms continues to apply.

Recovering the equilibrium allocation and the production function of matches between firms and workers is made difficult by the fact that only wages are observed, but actual matches' productivity is not. Abowd, Kramarz, and Margolis (1999) first studied this distributional assignment using a regression framework with firm and worker fixed effects which suggested very little sorting of more productive workers into more productive firms. Eeckhout and Kircher (2011) shows that, if the sign sorting is not always recoverable from wage data only, the strength of sorting is. Hagedorn, Law, and Manovskii (2012) show how information about co-workers can be used to identify non-parametrically the allocation distribution and the production function. Their approach utilizes the fact that with Nash bargaining, the wage ranks workers by type within each firm. They propose a rank aggregation method and demonstrate that it performs well on realistic sample size. Yet in their framework identical workers are paid similar wages when employed in the same firm.

In this paper we develop an equilibrium search model with two-sided heterogeneity, on-the-job search, training cost and vacancy creation. In the model, firms invest to create new positions, but each position can only hire exactly one worker. This creates a capacity constraint that generates sorting in equilibrium: firms and workers can decide not to match if it is more profitable to continue searching. This differs from Bagger and Lentz (2008) where sorting happens because higher ability workers move more quickly to better firms.

In section 1 we introduce the model formally, define the equilibrium and characterize some monotonicity properties. In section 2 we develop a constructive proof for the non parametric identification of the allocation of workers to firms and the underlying production function.

2.2 The model

We consider an economy populated by fixed numbers of workers and firms, all infinitely lived, risk-neutral, and discounting the future at rate r. Time is continuous.

2.2.1 Agents, technology and preferences

The economy is composed of a continuum of workers index by x and a continuum of firms index by (ϵ, y). The index x captures the ability of a worker, y represents the productivity of a firm and ϵ, the ability a firm has to create new positions. More formally, a firm (y, ϵ) can create n new jobs per period at convex cost $c(n, y, \epsilon)$ and a worker x employed in firm y produces output $f(x, y)$ every period where $f_x > 0, f_y > 0$. Employed workers are paid a wage $w \in \mathbb{R}$ which depends on their employment history. A firm employs multiple workers and we denote by $h(x, w|y, \epsilon)$ the mass of workers employed by a firm of type (y, ϵ) assuming firms to be large. Jobs are destroyed at exogenous rate δ. We denote $v(y, \epsilon)$ the number of unfilled jobs for that firm. Jobs are unfilled when originally created and when workers decide to move to a different firm.

Workers and firms are risk neutral, forward looking and discount at rate r. The mass of unemployed worker is denoted $u(x)$.

2.2.2 Meeting technology and matching decision

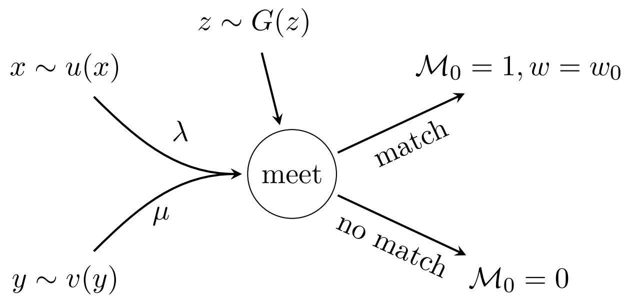


Figure 2.1: Meeting process and matching decision for unemployed

Workers and vacant jobs are brought together through a random meeting technology characterized by λ, κ and $\mu := \lambda \int u / \int v$. With probability $\kappa\lambda$ (or λ) an (un)employed worker meets a vacancy randomly drawn from $v(y,\epsilon)$. Conversely with probably $\kappa\mu$ a given vacancy meets an employed worker randomly drawn from $h(x,w,y,\epsilon)$ and with probability μ a random unemployed worker from $u(x)$. When a worker and firm meet, a positive random training cost $z \sim G(z)$ is drawn and must be paid if the match is formed.

Then comes the matching decision. When a firm (y,ϵ) meets an unemployed worker x with training cost z, the worker and the firm enter a bargaining process over the wage that splits the generated surplus. The outside option of the worker is remaining unemployed and the firm's is the present value of vacancy. We denote by $\mathcal{M}_0(x,y,\epsilon,z) \in \{0,1\}$ the matching decision and by $w_0(x,y,\epsilon,z)$ the outcome wage. The wage w_0 is set by generalized Nash bargaining.

In the case where a vacancy (y,ϵ) is matched to an already employed worker x in firm (y',ϵ') at some wage w', the two firms enter a Bertrand competition over the wage offered the worker. We call $\mathcal{M}_1(x,y,\epsilon,y',\epsilon',z') \in \{0,1\}$ the matching decision. It is equal to one if the worker is poached by the new vacancy. We also denote by $w_1(x, y, \epsilon, y', \epsilon', z')$ the outcome wage. $\mathcal{M}_1$ and w_1 are pinned down in equilibrium by the Bertrand competition.

2.2.3 Within period timing

First is production where employed worker collect their wage w, unemployed workers collect their flow value b, firms collect output $f(x, y)$ for each of their matches, choose how many new vacancies $n(y, \epsilon)$ to create and pay the vacancy creation cost $c(n, y, \epsilon)$. Second is the meeting stage. Vacancies find unemployed (employed) workers randomly at rate μ ($\kappa\mu$), workers find vacancies at rate λ ($\kappa\lambda$) and for each meeting the training cost z is drawn. Third is the matching decision and wage determination. Workers from unemployment bargain with their vacancies with outcome $(\mathcal{M}_0, w_0)$. Employed workers with an outside offer received the Bertrand outcome $(\mathcal{M}_1, w_1)$.

2.2.4 Flow equations for distributions

$(\mathcal{M}_0, w_0, \mathcal{M}_1, w_1)$ and $\lambda, \kappa, \delta, G$ define a Markovian law of motion for the distributions (h, u, v). We start by writing the flow equation for the distribution of unemployed workers

$$u_{t+1}(x) = u_t(x)\left[1 - \lambda \iiint \mathcal{M}_0(x, y, \epsilon, z) g(z) v(y, \epsilon)\, \mathrm{d}y\, \mathrm{d}z\, \mathrm{d}\epsilon\right] + (1 - u_t(x))\delta,$$

and we then consider the flow equation for the distribution of vacancies. This is given by

$$\begin{aligned} v_{t+1}(y, \epsilon) = n_t(y, \epsilon) + v_t(y, \epsilon)\Big[1 - \mu \iint \mathcal{M}_0(x, y, \epsilon, z) g(z) u(x)\, \mathrm{d}x\, \mathrm{d}z \\ - \kappa\mu \iiiint \mathcal{M}_1(x, y, \epsilon, y', \epsilon', z) G(z) h(x, w, y', \epsilon')\, \mathrm{d}y'\, \mathrm{d}z'\, \mathrm{d}\epsilon'\, \mathrm{d}x\Big], \end{aligned}$$

and finally the flow equation for the joint distribution of workers and firms is given in Appendix. We call Γ_1 the implied Markov transition kernel for the worker over the state space $\mathbb{X}, , \mathbb{R}$.

2.2.5 Equilibrium Definition

[SSE] Given primitives $f, G, r, \beta, \lambda, \kappa, c, b, \delta$, a **Stationary Search Equilibrium** ξ is defined by distributions h, u, v, matching and wages outcomes $\mathcal{M}_0, w_0, \mathcal{M}_1, w_1$ and firm job creation decision $n(y, \epsilon)$ such that:

(i) $(\mathcal{M}_0, w_0)$ solves generalized Nash bargaining between unemployed workers and vacancies (2.2.2) taking (h, u, v) as given

(ii) $(\mathcal{M}_1, w_1)$ solves the Bertrand competition (2.2.2) taking (h, u, v) as given

(iii) (h, u, v) are the stationary distributions of the Markovian law of motions generated by $(\mathcal{M}_0, w_0, \mathcal{M}_1, w_1)$ and $\lambda, \kappa, \delta, G$.

(iv) firms optimally choose $n(y, \epsilon)$

It is well known that the equilibrium might not be unique, and that for a set of primitives $f, G, r, \beta, \lambda, \kappa, c, b, \delta$ their might exists several such ξ. In the identification part of this paper, we will show that both ξ and the primitives are identified. The procedure itself identifies the realized equilibrium among multiple possible ones. When generating counterfactual, we should then worry about checking for equilibrium multiplicity, but during estimation, we do not need to worry about it.

2.3 Properties of the equilibrium

Let $\mathcal{U}(x)$ denote the value of unemployment and let $\mathcal{V}(y)$ be the value of a vacancy for a job of type y. Let $\mathcal{P}(x, y)$ denote the value of all future incomes that the worker and the job are going to generate. We define the match surplus as

$$\mathcal{S}(x, y) := \mathcal{P}(x, y) - \mathcal{U}(x) - \mathcal{V}(y),$$

and we define $\omega(x, y, s)$ the wage that delivers a surplus s to the w.

When an unemployed worker of type x meets a vacancy of type y, a non-negative adjustment cost (training cost, mobility cost, etc.) is drawn from a

Figure 2.2: Worker meets outside offer

distribution G. This adjustment cost is sunk. Then, if $\mathcal{S}(x,y) \geq z$, the worker is hired according to a contract specifying a fixed wage w. This means that a worker's wage can be modified only if both parties agree to renegotiate. This happens only if one of the two parties has a credible threat to break the match.

We assume that the value of the wage contract negotiated with an unemployed worker is equal to the value of unemployment plus a share β of the total surplus net of the adjustment cost, $\mathcal{S}(x,y) - z$. That is a wage $w_0(x,y,z)$ such that

$$\mathcal{W}\left(w_0(x,y,z),x,y\right) = \mathcal{U}(x) + \beta\left[\mathcal{S}(x,y) - z\right],$$

where $\mathcal{W}(w,x,y)$ denotes the present value of a wage contract w. The employer receives the value $\mathcal{V}(y) + (1-\beta)\left[\mathcal{S}(x,y) - z\right]$ if it pays $w_0(x,y,z)$.

As shown in Figure 2.2, when an employed worker of type (x,y) currently receiving a value $\mathcal{W} = \mathcal{W}(w,x,y)$, with $0 \leq \mathcal{W} - \mathcal{U}(x) \leq \mathcal{S}(x,y)$, meets a vacancy of type y' with training cost z', the two firms enter a Bertrand competition and three things can happen.

First if $S(x,y') - z' > S(x,z)$ the worker moves to the new firm and extract the surplus from y, she gets $\mathcal{W} = \mathcal{S}(x,y) + \mathcal{U}(x)$. Otherwise if $S(x,z) > S(x,y') - z' > \mathcal{W} - \mathcal{U}(x)$, the worker stays with firm y and extract a wage increase thanks to the outside offer. She gets $\mathcal{W} = S(x,y') - z' + \mathcal{U}(x)$. Finally if $\mathcal{W} - \mathcal{U}(x) > S(x,y') - z'$ nothing happens, the value and the wage do not change.

The present value of unemployment. Consider a worker of type x who is unemployed for a whole period. During that period she earns a flow income of $b(x)$ and at the end of the period she meets with a vacancy y with probability $\lambda v(y)$. The match is consummated if $\mathcal{S}(x,y) \geq z$, and the workers receives a share $\beta\left[\mathcal{S}(x,y)-z\right]$ of the surplus. Hence

$$r\mathcal{U}(x)=b(x)+\beta\int\mathcal{G}\left[\mathcal{S}(x,y)\right]\lambda v(y)\,\mathrm{d}y, \tag{2.1}$$

where we denote $\mathcal{G}(t)\equiv\int\max\{t-z,0\}g(z)\,\mathrm{d}z$.

The present value of a vacancy. Let $\mathcal{V}(y)$ denote the expected profit for a vacant job

$$r\mathcal{V}(y)=(1-\beta)\int\mathcal{G}\left[\mathcal{S}(x,y)\right]\mu u(x)\,\mathrm{d}x \\ +\iint\mathcal{G}\left[\mathcal{S}(x,y)-\mathcal{S}(x,y')\right]\kappa\mu h(x,y')\,\mathrm{d}x\,\mathrm{d}y'. \tag{2.2}$$

A firm operates only if the value of a vacancy is positive. To avoid the issue that $\mathcal{V}(y)$ might be negative for some y, we shall assume zero cost of maintaining vacancies, $\gamma=0$.

The match surplus. We now turn to the total surplus $\mathcal{S}(x,y)$ of a match a match (x,y). Flow output is $f(x,y)$. The match faces a flow probability of dissolution δ, in which case the continuation value is 0 for the firm (a job is destroyed), and $\mathcal{U}(x)$ for the worker. Otherwise, with flow probability $\kappa\lambda v(y)G\left(\mathcal{S}(x,y')-\mathcal{S}(x,y)\right)$, the worker is poached by a vacancy of type y'. In this case, the worker moves and pockets a value of $\mathcal{U}(x)+\mathcal{S}(x,y)$, while the firm is left with a vacant job worth $\mathcal{V}(y)$, so that the net continuation *gain* for the firm-worker match is 0. Summing up, match value solves:

$$(r+\delta)\mathcal{S}(x,y)=f(x,y)-r\mathcal{U}(x)-(r+\delta)\mathcal{V}(y). \tag{2.3}$$

The present value of being employed. Denoting $\mathcal{W}(w,x,y)$ for the value of a wage w given match characteristics (x,y), we have

$$(r+\delta+\kappa\lambda)\mathcal{W}(w,x,y) = w + \delta\mathcal{U}(x)$$
$$+\kappa\lambda \iint \max\Big\{\mathcal{U}(x)+\min\left\{\mathcal{S}(x,y),\mathcal{S}(x,y')-z'\right\},\mathcal{W}(w,x,y)\Big\} g(z')v(y')\,\mathrm{d}z'\,\mathrm{d}y',$$

as a poached worker gets the second price $\mathcal{U}(x)+\min\left\{\mathcal{S}(x,y),\mathcal{S}(x,y')-z'\right\}$ as long as it is greater than her reservation value $\mathcal{W}(w,x,y)$.

Job Creation.

A firm (y,ϵ) controls its stock of available jobs, and compensates for job obsolescence, by creating $n(y,\epsilon)$ new jobs per period at a cost $c(n,y,\epsilon)$:

$$n(y,\epsilon) = \arg\max_n n\mathcal{V}(y) - c(n,y,\epsilon),$$

where $\mathcal{V}(y)$ is the present value of a vacant job and $c(n,\epsilon)$ is the cost of investments in tools, computers, etc, that accompany the creation of n jobs. Hence,

$$c'[n(y,\epsilon),y,\epsilon] = \mathcal{V}(y). \tag{2.4}$$

Wages

The preceding Bellman equation readily defines w as a function of (x,y) and the surplus to the worker $\mathcal{W}(w,x,y)-\mathcal{U}(x)$. For any feasible value $s\in[0,\mathcal{S}(x,y)]$ of the worker surplus, the contract stipulates a wage

$$\begin{aligned}\mathrm{w}(s,x,y) &= r\mathcal{U}(x) + (r+\delta)s\\ &\quad -\kappa\lambda\int\Big\{\mathcal{G}\left[\mathcal{S}(x,y')-s\right]-\mathcal{G}\left[\mathcal{S}(x,y')-\mathcal{S}(x,y)\right]\Big\}v(y')\,\mathrm{d}y'.\end{aligned} \tag{2.5}$$

It thus follows that, for $\mathcal{S}(x,y)-z\geq 0$, the wage coming out of unemployment for a worker x into firm (y,z) is given by $\mathrm{w}\left(\beta\left[\mathcal{S}(x,y)-z\right],x,y\right)$. Then for a worker x transitioning from a firm y' to a firm (y',z') the wage is given by $\mathrm{w}\left(\mathcal{S}(x,y')-z,x,y\right)$. Notice that the maximal wage that a worker x can be offered for a job y is when she gets the whole surplus, i.e. $s=\mathcal{S}(x,y)$:

$$\begin{aligned}\overline{\mathrm{w}}(x,y) &:= \mathrm{w}\left[\mathcal{S}(x,y),x,y\right] \\ &= r\mathcal{U}(x)+(r+\delta)\mathcal{S}(x,y)=f(x,y)-(r+\delta)\mathcal{V}(y).\end{aligned} \tag{2.6}$$

Similarly, the minimal wage for a worker x at a job y is obtained for $s=0$, which happens when the adjustment cost z is equal to the whole surplus $\mathcal{S}(x,y)$, and the contract value is equal to the value of unemployment $\mathcal{U}(x)$:

$$\begin{aligned}\underline{\mathrm{w}}(x,y) &:= \mathrm{w}(0,x,y) \\ &= r\mathcal{U}(x)-\kappa\lambda\int\Big\{\mathcal{G}\left[\mathcal{S}(x,y')\right]-\mathcal{G}\left[\mathcal{S}(x,y')-\mathcal{S}(x,y)\right]\Big\}v(y')\,\mathrm{d}y'.\end{aligned} \tag{2.7}$$

This minimum wage is attained for all (x,y) if distribution G has a large support, for example $\mathbb{R}_+$. We assume that functions $f(x,y)$ and $b(x)$ are bounded and twice continuously differentiable, and we omit the proof that this property is passed on to values $\mathcal{U}(x),\mathcal{V}(y)$ and $\mathcal{S}(x,y)$, and to the wage function $\underline{\mathrm{w}}(s,x,y)$.

Lemma 7. *If $f(x,y)$ is increasing in x and y, $b(x)$ is nondecreasing, and c is convex, then the following monotonicity properties hold true:*

(i) *$\mathcal{U}(x)$ is increasing in x.*

(ii) *$\mathcal{V}(y)$ is increasing in y.*

(iii) *The maximal wage $\overline{\mathrm{w}}(x,y)$ is increasing in x and so is $\overline{\overline{\mathrm{w}}}(x)\equiv\max_y\overline{\mathrm{w}}(x,y)$.*

(iv) *If in addition f is supermodular ($\frac{\partial^2 f(x,y)}{\partial x\partial y}>0$), then $\mathcal{S}(x,y)$ and $\overline{\mathrm{w}}(x,y)$ are also supermodular and the probability of moving from y to $y'>y$ increases with x.*

Proof. See Appendix 2.B. □

2.4 Identification

In this section we address the issue of the identification of the model given linked employer-employee data. Here we think of the data as a collection of N independent worker trajectories drawn from the stochastic process $\Gamma_t = (X, E_t, R_t, J_t, Y_t, C_t), t \geq 1$, defined with respect to a filtered probability space $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, \mathbb{P})$, where E_t is the employment status ($E_t = 1$ if currently hired and 0 if unemployed), R_t stands for a worker's wage at time t (missing if the worker is unemployed) and J_t is the identifier of the firm employing the worker (if the worker is currently employed; missing otherwise), Y_t is the type of the firm J_t, C_t its vacancy creation cost and X is the type of the worker. (E_t, R_t, J_t, I_t) is observed and (X, Y_t, C_t) is not. Secondly we are also interested in computing information at the firm level. For instance we might want to compute the average wage in the firm. Given the set of firm identifier $\mathbb{J}$ we can properly define the conditional probability with respect to any firm $J \in \mathbb{J}$.

We then want to express sufficient conditions for a given process $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, \mathbb{P})$ to be generated from a particular model ξ . This ties the probability measure $\mathbb{P}$ to the properties the endogenous distributions and decision rules of the model: The random process Γ_t defined on $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, \mathbb{P})$ is generated by the equilibrium ξ if

(i) $X \sim \mathbb{U}([0,1])$

(ii) $\mathbb{P}\{X \leq x | E_t = 0\} = \int_0^x u(x)\,\mathrm{d}x$

(iii) $\mathbb{P}\{X \leq x, Y_t \leq y, R_t \leq w, C_t \leq \epsilon | E_0 = 1\} = \int_0^x \int_0^y h(x, y, w, \epsilon)/u(x)\,\mathrm{d}x\,\mathrm{d}y$

(iv) $\Gamma_{t+1}|\Gamma_t$ is implied by $(\mathcal{M}_0, w_0, \mathcal{M}_1, w_1)$ and $\mu, \lambda, \kappa, \delta, G$.

(v) There exists two deterministic functions $j^*$and ϵ^* in $\mathbb{J} \to [0,1]$ such that $\forall \omega, t,\ Y_t = j^*(J_t), C_t = \epsilon^*(J_t)$

Property (i, ii,iii) insure that the distribution in $\mathbb{P}$ match the distributions implied by the model. In includes the link between the job creation cost and the firm size. Property (iv) guarantees that law of motion of the random process generated on the filtration is distributed according to the law of motions implied by the equilibrium on the model. It guarantees that wages and mobility are correctly generated. Finally property (v) makes sure that the firm type does not change over time and across observations. The identification exercise is to use those properties without using (X, Y) to reconstruct the model. We now proceed to show that this information allows to identify the structural parameters under the following assumption.

Assumption 1.

(i) *Let* $z^* = \max_{x,y} \mathcal{S}(x, y)$. $G(0) = 0$ *and* G *has continuous support on* $[0, z^*]$

(ii) $f(x, y)$ *is differentiable and* $f_x > 0, f_y > 0$

(iii) $c(n, y, \epsilon)$ *is differentiable, convex in* n, $c(0, y, \epsilon) = 0$, *increasing in* $\epsilon \in [0, 1]$

Condition (i) implies that the support of G is large enough so that, for all x, the set of couples (y, z) such that $\mathcal{S}(x, y) = z$ is not empty. Such match combinations (x, y, z) are only marginally profitable. Similarly we assume that $G(z)$ is positive everywhere on $[0, z^*]$ to guaranty that with some positive probability every match with positive surplus will be formed. Condition (ii) imposes comparative advantage in production. Finally (iii) restricts the intercept of the cost function.

We can now state the main result of the paper.

Theorem 1. ***Identification.*** *Set the discount rate* r. *Under Assumption 1, knowing conditional expectations on* $(\Omega, \mathcal{A}, \{\mathcal{A}_t\}_{t\geq 0}, \mathbb{P})$ *for all observables, generated from an equilibrium* ξ, *and knowing the aggregate measure of vacancies* V, *then all primitives* f,G,β,λ,κ,c,b, δ *and endogenous values of* ξ *are identified.*

Proof. The proof of Proposition 1 is constructive and a direct implication of Lemma 1-9. □

The proposition states that knowing the random process Γ_t on observables only is enough to measure directly the primitives of the model.

Lemma 8. *The worker type is identified:*

$$\forall \omega,\ \overline{R} = \mathcal{Q}_{\bar{R}}(X).$$

where $\overline{R} := \max_t \{R_t : E_t = 1\}$ *is the* $\mathcal{A}$*-measurable maximum wage for each* ω*, and* $Q_{\overline{R}}(\cdot)$ *its quantile function.*

Proof. By Proposition 7(iv), the highest wage that a worker will attain in her career, $\overline{\overline{\mathrm{w}}}(x) = \max_y \overline{\mathrm{w}}(x, y)$, is increasing in x. By construction, the $\mathcal{A}$-measurable variable $\overline{R} = \overline{\overline{\mathrm{w}}}(X)$ is a deterministic and monotonic function of the random variable X. □

Lemma 9. $\mathcal{U}(x)$ *is given by:*

$$\forall x \quad \mathcal{U}(x) = \mathbb{E}_J \mathbb{E}_t \left[W_t \mid X = x, E_t > E_{t-1}, J_t = J, R_t = R_{min}(x, J_t) \right],$$

where $W_t := \sum_{\tau=t}^{\infty} \frac{R_\tau}{(1+r)^\tau}$ *is the* $\mathcal{A}$*-measurable realized present value for each* (ω, t)*, and* $R_{min}(x, J) := \min_{\omega \in \Omega, t \in T} \{R_t : E_t > E_{t-1}, J_t = J, X = x\}$ *is the minimum wage paid to worker* x *in firm* J*.*

Proof. It follows from Subsection 2.3 that the minimum wage ever paid to a worker x by a firm y is $\underline{\mathrm{w}}(x, y) = \mathrm{w}(0, x, y)$, delivering the same present value as unemployment. This is just stating that the worse contract a worker would accept is a contract that makes her indifferent to remaining unemployed. However at this point we can't aggregate over different firms since Y is latent. We can do this averaging inside each firm, and then aggregate.

For each firm in the data, we can define the smallest wage collected by any worker of a given type x. This requires us to take the expectation across workers that worked at least once in a firm J. We can then average across all firms. The only thing required at this point is that enough x-type workers visited firm J, but this will always be the case as time grows. Then $\forall \omega \in$

$\Omega, x \in \mathbb{X}, t \in T$ $\underline{R}(x, J_t) = \underline{\mathrm{w}}(x, Y_t)$ as long as workers with type x do work in firm J. We can now construct the value of unemployment using observables from the data. Conditional on collecting R_{min}, the expected present value to the worker delivers exactly $\mathcal{U}(x)$ and gives the expression stated in the Lemma. □

Further more, since once the worker becomes unemployed again, he will be getting $\mathcal{U}(x)$, one can use a finite sum to extract the value of unemployment. The function $\mathcal{U}(x)$ is identified if, for all x provided that that given x works at least in one firm, and that the z shock is large enough to deliver the reservation wage. The latter will always be true under Assumption 1.

Lemma 10. *Bargaining power β is given by:*

$$\beta = \mathbb{E}_J \left[\frac{\mathbb{E}_0 \left[W_0 | E_1 > E_0, X = x, R_1 = R_{max}(x, J_1), J_1 = J \right] - \mathcal{U}(x)}{S(x, J)} \right],$$

where the surplus at firm J is given by

$$S(x, J) = \mathbb{E}_t \left[W_t | J_t \neq J_{t-1}, E_t = E_{t-1} = 1, X = x, J_{t-1} = J \right] - \mathcal{U}(x).$$

Proof. We are going to use a similar procedure here. The highest wage collected on coming out of unemployment delivers the highest possible surplus. The present value associated with that wage delivers $\beta S(x, y)$ to the worker. Secondly workers going through a job to job transition actually extracts $S(x, y)$ from their current firm. More precisely:

$$\begin{aligned} \forall \omega, t \qquad & \mathbb{E}_t \left[W_t | J_t \neq J_{t-1}, E_t = E_{r-1} = 1, X = x \right] = S(x, J_t) + \mathcal{U}(x) \\ & \mathbb{E}_t \left[W_t | E_t > E_{t-1}, X = x, R_t = \overline{R}(x, J_t) \right] = \beta S(x, J_1) + \mathcal{U}(x) \end{aligned}$$

selecting workers entering a given firm J with the highest entry wage, and following them through their next job and then for ever. □

Here however because the y value must match at the numerator and denominator, we need to aggregate over the J outside the fraction.

Lemma 11. *G is given by*

$$\forall z \in [0, z^*] \quad G(z) = \mathbb{P}_{J,w}\left\{ \frac{\mathbb{E}\left[W_1 | X = x, E_1 > E_0, J_t = J, R_1 = w\right] - \mathcal{U}(x)}{\beta} > z \right\}.$$

Proof. Conditional on $X = x$ we already have a measurement of $S(x, J)$ for any J using the subsequent job-to-job transition. Using the worker surplus conditional on entry wage we can compute the distribution $\beta\left(S(x, Y) + z\right)$ and as such we can estimate the distribution over z.

$$\forall x, J, Z_1 \; \mathbb{E}\left[W_1 | X = x, E_1 > E_0, J_t = J, R_1 = w\right] = \beta(S(x, j^*(J)) + w_0^{-1}(w|x, j^*(J))) + \mathcal{U}(x)$$

where the $w_0^{-1}(w|x, y)$ is the z from the model that delivers wage w for that particular (x, y):

$$w_0^{-1}(w|x, y) := \{z | w_0(x, y, z) = w\}$$

This means that the distribution of z is directly identified. G is only recovered on part of its support. We assume that G is analytic, and so that its global properties can recovered from local identification. □

Lemma 12. *The efficiency of on-the-job search κ is given by*

$$\kappa = \frac{\mathbb{P}\{R_t > R_{t-1} \cup J_t \neq J_{t-1} | X, R_{t-1} = R_{min}(J, X)\}}{\mathbb{P}\{E_t > E_{t-1} | X\}},$$

and the separation rate is given by

$$\delta = \frac{\mathbb{P}\{E_{t+1} < E_t\}}{\mathbb{P}\{E_t = 1\}}.$$

Intuitively, when the worker is collecting the lowest wage, any viable match will increase his current wage or trigger a job transition since he is currently getting his reservation value. The denominator controls for the fact that not all meeting are with viable partners.

Lemma 13. *The firm type is identified by the rank of* $\hat{\mathcal{V}}(J)$ *where*

$$\hat{\mathcal{V}}(J) = (1-\beta) \int \mathcal{G}\left[\mathcal{S}(x,J)\right] u(x)\,\mathrm{d}x + \kappa \iint \mathcal{G}\left[\mathcal{S}(x,J) - S\right] \gamma(S,x)\,\mathrm{d}x\,\mathrm{d}S.$$

and $\Gamma_{Sx} = \int \gamma$ *is joint distribution of* (S,x) *in the population*

$$\Gamma(S,x) = \mathbb{P}\{X \le x \cup \mathcal{S}(X,J) \le S\}.$$

Proof. We know from Proposition 7 that the value of the vacancy is increasing in y. We only need to show that $\hat{\mathcal{V}}$ is a monotonic transformation of $\mathcal{V}(y)$. Recall its expression:

$$r\mathcal{V}(y) = (1-\beta)\mu \int \mathcal{G}\left[\mathcal{S}(x,y)\right] u(x)\,\mathrm{d}x \\ + \kappa\mu \iint \mathcal{G}\left[\mathcal{S}(x,y) - \mathcal{S}(x,y')\right] h(x,y')\,\mathrm{d}x\,\mathrm{d}y'. \quad (2.8)$$

We can ignore μ for now since it is an affine transformation. From the previous section we have identified $S(x, j^*(J))$. Yet we still do not know the actual y value itself. However given that G is also identified, and so are β and $u(x)$, the first integral term of $\mathcal{V}(J)$ is also computable from the random process itself. The second integral requires the distribution of surplus the firm draws from in the population. However we do observe the joint distribution over $(S(x, j^*(J)), x)$ in the data since we were able to measure $S(x, j^*(J))$ and x from previous Lemma. Going from $\mathcal{V}(y)$ to $\hat{\mathcal{V}}(J)$ is just a change of variable from y' to $S = S(x, y')$. This allows us to construct $\hat{\mathcal{V}}(J)$ for each J, which is an affine transformation of $V(y)$ and so reveals Y_t and the $j^*(\cdot)$ function. □

Lemma 14. *The Surplus* $\mathcal{S}(x,y)$ *is identified by:*

$$\mathcal{S}(x,y) = \mathbb{E}_t\left[W_t | J_t \neq J_{t-1}, E_t = E_{t-1} = 1, X = x, Y_{t-1} = y\right] - \mathcal{U}(x),$$

and the steady state stock of vacancies by:

$$v(y) \propto \mathbb{P}\{E_t > E_{t-1} | X = x, J_t = J\} / G\left[\mathcal{S}(x, y)\right],$$

and where the normalizing constant and μ are pinned down by the aggregate mass of vacancy V.

Proof. This comes from the flow of workers out of unemployment $\mu v(y) u(x) G\left[\mathcal{S}(x, y)\right]$ and the fact that the probability space is generated by the model. Importantly, by calculating the meeting rate conditional on x, one removes the problem that some meeting do not become matches. We established previously that $S(x, J)$ was identified for any J. We have also shown that the actual value of Y is also identified for each J. So $S(x, y)$ is identified for each (x, y). □

It immediately follows that the production complementarity between x and y, as defined by the information contained in the second partial derivative $\frac{\partial^2 f(x,y)}{\partial x \partial y} = (r + \delta) \frac{\partial^2 \mathcal{S}(x,y)}{\partial x \partial y}$, is identified.

Lemma 15. *$b(x)$ is given by :*

$$b(x) = r\mathcal{U}(x) - \beta \int \lambda v(y) \mathcal{G}\left[\mathcal{S}(x, y)\right] \mathrm{d}y$$

and the match-production function $f(x, y)$ by

$$f(x, y) = (r + \delta)\mathcal{S}(x, y) - r\mathcal{U}(x) - (r + \delta)\mathcal{V}(y)$$

The Bellman equation for $\mathcal{U}(x)$ identifies $b(x)$ and consequently the production function $f(x, y)$ is non-parametrically identified.

Lemma 16. *The vacancy creation cost is given by*

$$\frac{\partial c}{\partial n}(\delta \mathcal{Q}_{l|y}(\epsilon | y), \epsilon) = \mathcal{V}(y)$$

and the assumption that $c(0, y, \epsilon) = 0$, *where* $\mathcal{Q}_{l|y}(\epsilon|y)$ *is the quantile function of firm size conditional on* y.

Proof. We reconstruct this from the joint distribution of vacancy types y and average firm size. The first order condition of the vacancy investment is

$$\frac{\partial c}{\partial n}(n, y, \epsilon) = \mathcal{V}(y),$$

which links the firm size to the value of the vacancy

$$\frac{\partial c}{\partial n}(\delta l(y), y, \epsilon) = \mathcal{V}(y)$$

so from the joint distribution of y and long term firm size, we recover the cost function c. It is actually easier to uncover through the quantile distribution. Normalizing ϵ to $[0, 1]$ we have that which together with the condition that $c(0)$ pins down the vacancy creation cost function. $\square$

2.5 Conclusion

In this paper we study the empirical content of matched employer-employee data. We demonstrate how a model with two-sided heterogeneity, on-the-job search and vacancy creation is non-parametrically identified provided that the data is large enough in both sample size and time length. We hope to have achieved three things in this paper. The first is to convince that there is a lot of information in knowing who works where at what wage over time. Indeed this is enough to recover unobserved heterogeneity of both workers and firms in an environment where linear fixed-effect regression deliver biased estimates. Second we demonstrate how the firm size distribution in our model can be utilized to learn the job creation process and costs faced by firms. Finally, on a more practical matter, we hope that this paper will help applied work that deal with smaller sample size with picking informative moments to estimate model using methods of moments. For instance it is notable how important job to job transitions and the wages collected after that are to measuring the surplus function.

Bibliography

Abowd, J., F. Kramarz, and D. Margolis (1999): "High wage workers and high wage firms," *Econometrica*, 67(2), 251–333.

Atakan, A. E. (2006): "Assortative matching with explicit search costs," *Econometrica*, 74(3), 667–680.

Bagger, J., and R. Lentz (2008): "An empirical model of wage dispersion with sorting," *University of Wisconsin mimeo.*

Becker, G. S. (1974): "A theory of marriage," *Journal of Political Economy*, pp. 299–351.

Eeckhout, J., and P. Kircher (2010): "Sorting and decentralized price competition," *Econometrica*, 78(2), 539–574.

Eeckhout, J., and P. Kircher (2011): "Identifying sorting in theory," *The Review of Economic Studies*, 78(3), 872–906.

Hagedorn, M., T. H. Law, and I. Manovskii (2012): "Identifying Equilibrium Models of Labor Market Sorting," *Working Paper.*

Mortensen, D. (2005): *Wage dispersion: why are similar workers paid differently.* The MIT Press.

Shimer, R., and L. Smith (2003): "Assortative matching and search," *Econometrica*, 68(2), 343–369.

2.A More equilibrium properties

2.A.1 Equilibrium Worker and Vacancy Distributions

We now derive the steady-state distributions of worker skills in unemployment and within firms, and the distribution of vacancies by firm type.

Steady-State Unemployment Distribution. The steady-state unemployment distribution $u(x)$ satisfies the flow-balance equation

$$\delta \iint h(x,y,\epsilon)\,\mathrm{d}y\,\mathrm{d}\epsilon = u(x) \iint \lambda v(y,\epsilon)\, G\left[\mathcal{S}(x,y)\right]\,\mathrm{d}y\,\mathrm{d}\epsilon,$$

The LHS is the flow of employed workers of type x who are laid off in a small time interval. The RHS is the flow of unemployed workers of type x who find a job. Now, making use of the accounting equality $u(x) = 1 - \int h(x,y)\,\mathrm{d}y$, we have, at the steady-state equilibrium,

$$\delta\,(1-u(x)) = u(x) \int \lambda v(y)\, G\left[\mathcal{S}(x,y)\right]\,\mathrm{d}y, \tag{2.9}$$

Steady-State Employment Distributions. The following flow-balance condition determines the steady-state distribution of employment $h(x,y,\epsilon)$:

$$\begin{aligned}&\left[\delta + \int \kappa\lambda\mu v(y')\, G\left[\mathcal{S}(x,y') - \mathcal{S}(x,y)\right]\,\mathrm{d}y'\right] h(x,y,\epsilon)\\ &= \mu v(y,\epsilon) G\left[\mathcal{S}(x,y)\right] u(x) + \kappa\lambda\mu v(y,\epsilon) \int G\left[\mathcal{S}(x,y) - \mathcal{S}(x,y')\right] h(x,y')\,\mathrm{d}y'.\end{aligned} \tag{2.10}$$

where we make use of the marginal $v(y) = \int v(y,\epsilon)\,\mathrm{d}\epsilon$.

Steady-State Distribution of Vacancies. The flow of new vacancies of type y has two components: one is made of the new jobs $n(y,\epsilon)$ created in order to compensate for exogenous job obsolescence,

$$n(y,\epsilon) = \delta \int h(x,y,\epsilon)\,\mathrm{d}x \equiv \delta\ell(y,\epsilon),$$

where $\ell(y,\epsilon) = \int h(x,y,\epsilon)\,\mathrm{d}x$ is the size of a firm of type (y,ϵ); another corresponds to the flow of vacancies resulting from poaching. At the steady-state equilibrium, the number of vacancies which find an employee must equate the number of new vacancies:

$$\begin{aligned} v(y,\epsilon)\left[\int \mu u(x)\, G\left[\mathcal{S}(x,y)\right]\,\mathrm{d}x + \kappa \iint G\left[\mathcal{S}(x,y)-\mathcal{S}(x,y')\right]\mu h(x,y')\,\mathrm{d}x\,\mathrm{d}y'\right] \\ = n(y,\epsilon) + \kappa \iint G\left[\mathcal{S}(x,y')-\mathcal{S}(x,y)\right]\mu v(y')h(x,y)\,\mathrm{d}y'\,\mathrm{d}x, \end{aligned} \tag{2.11}$$

where the last component of the RHS is the flow of employees of jobs y who get successfully poached by a firm y'.

2.B Proof for Proposition 1

(i) $\mathcal{U}(x)$ is increasing in x

Differentiating equation (2.1),

$$r\mathcal{U}'(x) = b'(x) + \beta \int \lambda v(y) G\left[\mathcal{S}(x,y)\right] \frac{\partial \mathcal{S}(x,y)}{\partial x}\,\mathrm{d}y,$$

where

$$\frac{\partial \mathcal{S}(x,y)}{\partial x} = \frac{1}{r+\delta}\frac{\partial f(x,y)}{\partial x} - \frac{r}{r+\delta}\mathcal{U}'(x),$$

and reordering terms, we have

$$r\mathcal{U}'(x) = \frac{b'(x) + \beta\lambda \int v(y) G\left[\mathcal{S}(x,y)\right] \frac{\partial f(x,y)}{\partial x}\,\mathrm{d}y}{r+\delta+\beta\lambda\int v(y)G\left[\mathcal{S}(x,y)\right]\,\mathrm{d}y},$$

which is positive under the assumption that $\frac{\partial f(x,y)}{\partial x} > 0$ and $b'(x) \geq 0$.

(ii) $\mathcal{V}(y)$ is increasing in y

Differentiating equation (2.2),

$$r\mathcal{V}'(y) = (1-\beta)\int G\left[\mathcal{S}(x,y)\right]\frac{\partial \mathcal{S}(x,y)}{\partial y}\mu u(x)\,\mathrm{d}x \\ + \kappa \iint G\left[\mathcal{S}(x,y)-\mathcal{S}(x,y')\right]\frac{\partial \mathcal{S}(x,y)}{\partial y}\mu h(x,y')\,\mathrm{d}x\,\mathrm{d}y',$$

where

$$\frac{\partial \mathcal{S}(x,y)}{\partial y} = \frac{1}{r+\delta}\frac{\partial f(x,y)}{\partial y} - \mathcal{V}'(y),$$

and reordering terms, we have

$$\Big[r + (1-\beta)\int G\left[\mathcal{S}(x,y)\right]\mu u(x)\,\mathrm{d}x \\ + \kappa \iint G\left[\mathcal{S}(x,y)-\mathcal{S}(x,y')\right]\mu h(x,y')\,\mathrm{d}x\,\mathrm{d}y'\Big](r+\delta)\mathcal{V}'(y) \\ = (1-\beta)\int G\left[\mathcal{S}(x,y)\right]\frac{\partial f(x,y)}{\partial y}\mu u(x)\,\mathrm{d}x \\ + \kappa \iint G\left[\mathcal{S}(x,y)-\mathcal{S}(x,y')\right]\frac{\partial f(x,y)}{\partial y}\mu h(x,y')\,\mathrm{d}x\,\mathrm{d}y',$$

which is positive under the assumption that $\frac{\partial f(x,y)}{\partial y} > 0$.

(iv) The maximal wage $\overline{\mathrm{w}}(x,y)$ is increasing in x

By definition

$$\overline{\mathrm{w}}(x,y) = f(x,y) - (r+\delta)\mathcal{V}(y).$$

Hence, $\frac{\partial \overline{\mathrm{w}}(x,y)}{\partial x} = \frac{\partial f(x,y)}{\partial x}$ is positive if $\frac{\partial f(x,y)}{\partial x} > 0$. The Envelope Theorem guarantees that this property passes on to $\max_y \overline{\mathrm{w}}(x,y)$.

(iv) If f supermodular ($\frac{\partial^2 f(x,y)}{\partial x \partial y} > 0$), then $\mathcal{S}(x, y)$ and $\overline{\mathrm{w}}(x, y)$ are also supermodular

From the previous definition $\frac{\partial^2 \overline{\mathrm{w}}(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial x \partial y} > 0$;$\overline{\mathrm{w}}(x, y)$ is supermodular. Hence, $(r + \delta)\mathcal{S}(x, y) = \overline{\mathrm{w}}(x, y) - r\mathcal{U}(x)$ is also supermodular. Therefore, the probability of moving from y to $y' > y$ given x, $\kappa\lambda v(y)G\left[\mathcal{S}(x, y') - \mathcal{S}(x, y)\right]$, is increasing with x.

Chapter 3

Productivity Shocks, Dynamic Contracts and Income Uncertainty

3.1 Introduction

What are the drivers behind the observed earnings and employment uncertainty faced by workers in the labour market? How is this uncertainty mitigated by contracts between workers and firms in equilibrium? How is this transmission mechanism affected by policies? To address these questions, I develop a framework where workers face uncertainty about both their productivity and ability to locate new job opportunities and where firms choose optimally how wages respond to shocks. Changes in aggregate, firm level and worker specific productivities affect the value of a worker to an employer. At the same time, employed workers cannot immediately switch firms when current productivity decreases and unemployed workers might require several periods to locate a job opportunity. I show theoretically that in equilibrium, firms offer contracts that smoothly track worker's productivity in his current match, while responding with different intensity to different sources of shocks. I estimate the model using match employer-employee data and find that firm shocks accounts for 20% of a worker's permanent income uncertainty and that only about a third of underlying productivity gets passed through into wages. Employment transitions to unemployment and other jobs (40%) and worker shocks (40%) are the main sources of uncertainties since those are not insured by the firm. This confirms that unemployment insurance plays an important role in providing insurance that cannot be insured by the wage contract since the firm is unable to insure the worker when the employment relationship ends. However if generous unemployment insurance reduces earning risk, it also affects the employment level and the total output of the economy.

Earnings and employment uncertainty have important implications for welfare. A large body of literature has studied both theoretically and empirically the nature of the income process and quantified how it translates into consumption and wealth inequalities[1]. However, the income process itself is the

[1](MaCurdy, 1982; Blundell, Pistaferri, and Preston, 2008; Attanasio and Pavoni, 2011; Low, Meghir, and Pistaferri, 2009)

observed part of the complex employment agreement that links a worker to a job. The mechanism that defines this agreement in equilibrium, how the payments are delivered over time and how they respond to underlying productivity shocks has long been of high interest to the literature both theoretically and empirically. Knight (1921) first pointed out that one of the roles of the firm is to insure workers against productivity shocks. Baily (1974) and Azariadis (1975) formalized the idea and showed theoretically that when firms can sign long-term contracts, they fully insure their work force and offer fixed wage contracts even in the presence of demand shocks. Yet empirically income processes feature growth and employment risks (Altonji, Smith, and Vidangos, 2009; Low, Meghir, and Pistaferri, 2009).

Empirical evidence for the transmission of firms' shocks to workers' wages is provided by Guiso, Pistaferri, and Schivardi (2005). Using employer-employee matched data from Italy, they estimate how permanent and transitory productivity shocks of firms enter the wage equation of continuing workers. They report that full insurance of firms permanent shocks is rejected by the data. Their paper, however, uses a sample of continuing workers and does not control directly for the selection of workers in and out of firms. If workers who suffer the most from a drop in firm performance are also the ones leaving the sample, the effect of firm shocks is underestimated. Roys (2011) uses French firm data to estimate a model with homogenous workers and firm adjustment costs. He finds that firm permanent shocks affect employment and transitory shocks affect wages. The result however might be driven by the assumption that wages are set according to Nash bargaining which means that they are continuously renegotiated.

Contract theory offers answers to the apparent failure of the first best allocation. Harris and Holmstrom (1982) show that in a competitive market without worker commitment firms continue to insure against downward risk but have to increase the wage whenever productivity increases in order to retain the worker. Thomas and Worrall (1988) introduce the idea of a shock to job productivity by developing a model where a match between a firm and worker enjoy rents that can vary over time. They derive the optimal contract in an environment where the outside option is exogenous and show that, in a way similar to Harris and Holmstrom (1982), the wage remains constant until either the firm's or the worker's participation constraint binds. Burdett and Coles (2003) and Shi (2009) characterize the optimal contract when outside offers come from competing firms and the worker's decision is private; firms offer wages that increase with tenure to retain workers even-though they are risk averse and would prefer flat wages. Menzio and Shi (2010) extends this equilibrium framework by reintroducing match shocks and aggregate fluctuations, but do not characterize the optimal contract. Schaal (2010) does characterize the contract in a similar model but with homogeneous risk neutral workers. **?** derives the optimal contract with two sided lack of commitment but without on-the-job search or any private action from the worker, wage changes when outside options bind. To my knowledge, the current paper is the first to characterize the long term optimal contract offered in equilibrium by firms in an economy with search frictions, on-the-job search, firm and worker shocks and risk averse workers.

This paper makes three contributions to the existing literature. First, I document new findings about the co-movement of wages among co-workers which suggests larger transmission of firm shocks to wages than previously reported. Second, I characterize the optimal contract offered by competing firms in a directed-search equilibrium. Third, I estimate and evaluate quantitatively the model using linked employer-employee data.

The model builds on the directed search equilibrium of Menzio and Shi (2009), which allows for stochastic heterogeneity of firms and workers as well as worker risk aversion. Workers can search for new positions when employed and when unemployed. When on the job, the search decision is not observed and outside offers are not contractible by the firm. Firms can commit to any history-contingent long-term contract but the worker can walk away at any time. This contract flexibility is crucial because picking a particular form of wage setting might impose a specific level of insurance between the firm and the worker, whereas here it is determined by profit maximization. Flows of workers into firms is modeled using directed search, when searching for a new job, workers observe all contract offers and choose one to apply to. Each contract has a queue associated with it and each worker chooses the queue that maximizes the product between the return of the contract and the probability of getting picked from that queue. Directed search is a very natural extension of the competitive labour market that directly generates all the endogenous movement of workers in, out and across firms[2].

I show that in equilibrium firms post contracts that can be represented by a target wage that corresponds to the certainty equivalent of the current match productivity. Wages below that target wage will increase and wages above will decrease. The optimal contract presented here shares features of both Burdett and Coles (2003) and Hopenhayn and Nicolini (1997): firms back-load wages to incentivize workers to search less when profits are positive and front-load wages when profits are negative to incentivize the worker to search for a better position. When the match experiences a negative shock, the firm does not want to layoff the worker right away and decides to insure her in way similar to an optimal unemployment insurance scheme.

The empirical strategy of this paper utilizes the property that the wage smoothly tracks the target wage, which is subject to both worker and firm productivity shocks. Assuming that shocks to the worker and shocks to the firm are independent and that co-workers share the same firm productivity, shocks to the firm should affect all workers, whereas idiosyncratic shocks should affect them in an uncorrelated way. Using the auto-covariance and co-variance of co-workers' wages, I can identify how much of the wage movement is due to the firm relative to the worker and estimate the productivity process of both.

In Section 1, I present auxiliary models that will be used for estimation.

[2]The pioneering work in directed search is due to (Montgomery, 1991; Peters, 1991; Moen, 1997; Shimer, 1996; Burdett, Shi, and Wright, 2001; Shimer, 2001).

This section also motivates the economic question with evidence of risk transmission at the firm level in the Swedish matched employer-employee data. In Section 2 I formally present the equilibrium search model and I characterize the optimal contract. In Section 3, I present the estimation strategy and I discuss the identification of the model. This section also reports the estimation results. In Section 4, I put the model to work to answer the question of how much of income uncertainty is due to worker shocks and how much is due to the firm. A summary of the notation and all proofs are in the Appendix.

3.2 Earnings dynamics and participation

3.2.1 Data

The employer-employee matched data from Sweden links three administrative data-sets: the employment data, the firm data and the benefits data that track workers who are currently unemployed. The sample runs from 1993 to 2007 and covers around 6 million individuals. The firm data covers around 100,000 firms in four industries. The sample only covers firms with more than 10 employees, which means that some workers covered in the data work in a firm for which we do not have an identifier. On the worker side, all self-employed are dropped from the original sample, as well as some specific industries such as fisheries and the financial sector. I first de-trend the data with time dummies to remove any non stationary effects. I select individuals under 50 years of age, and, for moments computed at the firm level, I limit the data to firms with at least 25 employees.

3.2.2 Wage growth for job-stayers

In order to give an intuitive interpretation to moments computed from the data I introduce the following statistical model for residual log earnings of

	HS dropout	HS grad	Some college
residual wage variation σ_w^2	0.1274	0.1159	0.2033
	(0.000208)	(0.000132)	(0.000338)
worker transitory σ_v^2	0.0128	0.0123	0.014
	(0.000158)	(0.000117)	(0.000179)
worker permanent σ_ξ^2	0.0198	0.0173	0.0193
	(0.000238)	(0.000161)	(0.000242)
co-worker permanent σ_δ^2	0.0012	0.00146	0.00174
	(3.83e-05)	(3.11e-05)	(4.93e-05)
shared by co-workers	6.07%	8.41%	8.97%
	(0.19)	(0.173)	(0.309)
equivalent lottery	±3.47%	±3.82%	±4.17%
	(0.0555)	(0.0407)	(0.0591)

Standard errors are computed using clustered resampling. Wage differences are taken year on year. The equivalent lottery represents fair lottery over a permanent wage raise or cut in percent that would be equivalent to the share of variance common to co-workers.

Table 3.1. Residual income variance

continuing workers :

$$
\begin{aligned}
w_{ijt} &= \beta Z_t + \tilde{w}_{ijt} + v_{ijt} \\
\tilde{w}_{ijt} &= \tilde{w}_{ijt-1} + \delta_{jt} + \xi_{ijt},
\end{aligned}
$$

where i is the individual, j is the firm and t is time. Z_t is a yearly dummy, $\tilde{w}_{ijt}$ is the permanent wage, ξ_{ijt} is an idiosyncratic permanent shock to the wage and δ_{jt} is a permanent shock shared by all the workers in firm j. Wage growth shared by co-worker should be thought of as a firm specific event. The model parameters can be estimated using simple moments from individual wage growth and average wage growth within a firm (See Appendix 3.A.1). I report the estimates for each education group in Table 3.1.

The value of σ_δ is of interest as it represents the risk that co-workers share. To understand its monetary value, it is useful to think of the equivalent lottery that delivers a permanent percentage wage raise or cut. For instance for college graduates, every year, co-workers in a firm face the same lottery draw that delivers with 50 percent chance a wage raise of 4.17 percent and with 50 percent chance a 4.17 percent wage cut. This wage growth lottery

	HS dropout	HS grad	Col grad
τ	0.0287***	0.0217***	0.0181***
	(0.000955)	(0.000643)	(0.000679)
equivalent lottery	±0.537%***	±0.453%***	±0.399%***
	(0.0179)	(0.0134)	(0.015)

Standard errors are computed using clustered resampling. Wage differences are taken year on year. The equivalent lottery represents fair lottery over a permanent wage raise or cut in percent that would be equivalent to the share of variance common to co-workers.

Table 3.2. Income variance and value added shocks

is permanent and consequently 4.17 percent is economically significant. This provides evidence that part of the wage growth uncertainty is shared at the firm level.

3.2.3 Wage growth and value added

Quantitatively, the numbers presented in the previous section are larger than the one reported previously in the literature that focused on the link between value added and wages such as Guiso, Pistaferri, and Schivardi (2005) and Roys (2011). I replicate here a procedure similar to those papers to compare the Swedish economy to the French and Italian ones. I consider a simple unit-root model for the log value added per worker. The innovation shock μ_{jt} is then linked to the shock of permanent income among co-workers δ_{jt} from the previous section by the parameter τ:

$$\begin{aligned} y_{jt} &= \beta X_t + \tilde{y}_{jt} + u_{jt} \\ \tilde{y}_{jt} &= \tilde{y}_{jt-1} + \mu_{jt} \\ \delta_{jt} &= \tau \mu_{jt} + \nu_{jt} \end{aligned}$$

Table 3.2 contains the estimates for τ for each eduction group as well as the equivalent lottery implied by the amount of log wage growth uncertainty explained by shocks to value added. As in Guiso, Pistaferri, and Schivardi (2005) we see that the link between value added and wages is significantly different from zero. This provides evidence against the hypothesis of full insurance of firm shocks. The magnitude of the transmission of value added shocks to worker is economically small and similar to the values reported previously in the literature. Tables 3.1 and 3.2 suggest that shocks to value added can only explain a small part of the risk co-workers jointly share at the firm level. Consequently, I will focus the empirical analysis on wages within the firms rather than on value added.

3.2.4 Worker transitions

Finally it is also of interest to measure how changes at the firm level affect transitions of workers to unemployment and to other firms. When a firm receives a bad productivity shock, transitioning to another firm is a good way to insure income. This is precisely why studying the impact of search friction on the provision of insurance is important. Table 3.3 reports a linear probability models of worker transitions to unemployment and to other firms. Regressors include the mean wage of the firm and the mean wage change in the firm.

First we see that the mean wage in the firm affects negatively both transitions. This suggests that better firms pay higher wages and keep their workers longer. Interestingly however the worker's wage affects positively the probability to change firms. This can be explained by the fact that higher wages are more difficult to increase to prevent the worker form leaving, or it could be that higher earners move more frequently. We also note that an increase in firm average wage while keeping worker's wage constant affects positively the mobility of the worker. The results from Table 3.3 tell us that the risk of job loss is affected by changes in the firm. Not only do firms that pay higher wage seem to retain their workers longer, it also seems that a change in the firm's average wage does affect the rate at which workers loose their job. This source of risk associated with job loss can't be captured by the log wage models presented in the previous section that only looked at continuing workers.

	to another firm			to unemployment		
	HS dropout	HS grad	Col grad	HS dropout	HS grad	Col grad
(Intercept)	0.0134***	0.0185***	0.0245***	0.0163***	0.0154***	0.0101***
	(0.000118)	(6.48e-05)	(0.000107)	(0.000124)	(5.74e-05)	(6.76e-05)
worker wage	0.0329***	0.0468***	0.028***	-0.0142***	-0.0033***	-0.00178***
	(0.000386)	(0.000219)	(0.00027)	(0.000408)	(0.000194)	(0.000171)
firm wage change	0.0162***	0.0165***	0.0035***	0.0186***	0.0208***	0.0202***
	(0.00122)	(0.000782)	(0.0011)	(0.00129)	(0.000693)	(0.000695)
firm wage	-0.032***	-0.0522***	-0.0281***	-0.0115***	-0.0222***	-0.0132***
	(0.000673)	(0.000422)	(0.000494)	(0.00071)	(0.000374)	(0.000312)
N	2,450,855	9,788,831	4,246,564	2,450,855	9,788,831	4,246,564

Table 3.3: Transition probabilities to unemployment and other jobs

The model introduced in the rest of the paper will account for this additional employment risk.

3.3 The model

I present here an equilibrium model with search frictions and private worker actions. The key feature of the model is to embed the bilateral relationship between the firm and the worker, with productivity uncertainty, inside a competitive search equilibrium where firms compete to attract and retain workers.

In this model, ex-ante identical firms compete by posting long-term contracts to attract heterogeneous workers. Employed and unemployed workers observe the menu of contracts offered in equilibrium and decide which one to apply to. This process forms sub-markets of workers applying to particular contracts and firms offering them. Within each queue the matching between firms and workers is random. When choosing which sub-market to participate in, both firms and workers take into account the value of matching and the probability of matching. This probability is driven by how many firms and workers participate in a particular sub-market.

When matched, the contract specifies the wage after each possible history of shocks for the firm and workers. Given his wage profile, the worker chooses which sub-market to visit while employed and chooses effort, which directly affects the probability the current match remains intact. Both of these actions are private and so unobserved by the firm. Firms take this into account and post contracts that incentivize the worker's action in an optimal way. This will mean that in some cases the wage will adjust downward albeit in a smooth way. I now formally introduce the model.

3.3.1 Environment

Agents and preferences

Time is discrete, indexed by t and continues for ever. The economy is composed of a discrete uniform distribution of infinitely lived workers with ability indexed by $x \in \mathbb{X} = \{x_1, x_2 ... x_{n_x}\}$. Workers want to maximize expected lifetime utility, $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(w_t) - c(e_t)\right)$ where utility of consumption $u : \mathbb{R} \to \mathbb{R}$ is increasing and concave and cost of effort $c : \mathbb{R} \to \mathbb{R}$ is increasing and convex with $c(0) = 0$. Worker's ability x changes over time according to Markov process $\Gamma_x(x_{t+1}|x_t)$. Unemployed workers receive flow value of unemployment $b(x)$. The other side of the market is composed of a uniform distribution of ex-ante identical firms with active jobs and vacancies. Vacancies live for one period and become active jobs if matched with a worker.

An active job is characterized by the current worker ability x and the current match quality z. The match quality z evolves with an innovation ι_t drawn at the firm level such that $z_{t+1} = g(z_t, \iota_t)$. ι_t is a firm level shock that affects all continuing workers' TFP. New hires all start with $z = 0$. The function $g(\cdot, \cdot)$ is assumed to generate a monotonic transition rule. Every period a match (x_t, z_t) has access to a technology that produces $f(x_t, z_t)$. Worker's effort e affects the probability that the technology continues to exists next period. This captures the idea that a negligent worker might loose a client or break the machine and cause the job to disappear. The firm cares about the total discounted expected profit of each created vacancy.

Firms here operate constant return to scale production functions and can be thought of as one worker per firm. However, empirically one cannot aggregate firms with the same output as the history of productivity shocks affects the distribution of workers. For instance whether or not a firm had a very bad shock in the last period will affect the current distribution of workers beyond the current productivity. To pin down the distribution of workers in a given firm one needs to know the entire history of shocks.

Search markets

The meeting process between workers and firms vacancies is constrained by search frictions. The labour market that matches workers to vacancies is organized in a set of queues indexed by $(x, v) \in \mathbb{X} \times \mathbb{V}$ where x is the type of the worker and v is the value promised to her in that given queue. Firms can choose in which (x, v) lines they want to open vacancies and workers can choose in which v line associated with their type x they want to queue[3]. Each visited sub-market is characterized by it's tightness represented by the function $\theta : \mathbb{X} \times \mathbb{V} \to \mathbb{R}_+$ which is the ratio of number of vacancies to workers. The tightness captures the fact that a high ratio of vacancies to workers will make it harder for firms to hire. In a directed search model like the one presented here, the tightness is queue specific which means that different worker types could be finding jobs at different rates. In queue (x, v) a worker of type x matches with probability $p(\theta(x, v))$ and receives utility v. Firms post vacancies at unit cost η and when posting in market (x, v) the vacancy is filled with probability $q(\theta(x, v))$. $\phi(x, v)$ will denote the mass of vacancies created in market (x, v).

States and actions

A worker is either employed or unemployed and enters each period with a given ability x. When unemployed she collects benefit $b(x)$ and can search every period. When searching she chooses which sub-market (x, v) to visit, in which case she gets matched with probability $p(\theta(x, v))$ and if matched joins a job and receives lifetime utility v.

[3]Menzio and Shi (2009) Theorem 3 tells us that workers will separate by type in equilibrium if markets are indexed by the value that each type x would get in a particular sub-market ($\mathbf{v} = (v(x_1), v(x_2)...v(x_{n_x})) \in \mathbb{R}^{n_x}$,) and workers can apply to any. At equilibrium only a given type x visits a particular market. This market can then be represented directly by (x, v) as done in the current paper.

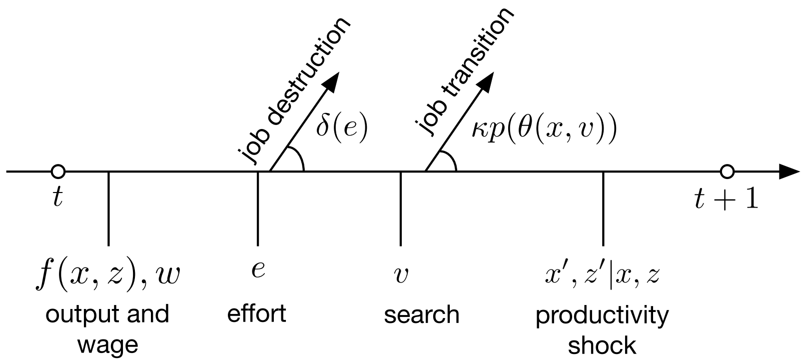


Figure 3.1: within period time line

An employed worker is part of a match and starts the period with a given ability level x and a current match quality z. The period is then divided in four stages as illustrated in Figure 3.1, first is production, the firm collects output $f(x,z)$ and pays the wage w to the worker. The worker cannot save, consumes all of w, chooses effort e and gets flow utility $u(w) - c(e)$. With probability $(1 - \delta(e))$, where $\delta(e)$ is decreasing in e, the employment persists to the next period. With probability $\delta(e)$ the worker moves to unemployment. In the search stage, the worker is allowed to search with efficiency κ. When searching she chooses which sub-market (x,v) to visit and gets matched with probability $\kappa p(\theta(x,v))$. If matched she moves to a new match where she will enjoy v and the current job will be destroyed. If the worker is not matched to a new job, the current job persists, a new x' is drawn conditional on the old one, and a firm level shock ι is drawn to update z. In summary, in every period an active job chooses the wage w, and the worker chooses effort e and which sub-market (x,v) to search in. Because $c(0) = 0$ the worker can quit in every period if the firm does not promise enough. By choosing v and e the worker controls his transition to other jobs and to unemployment.

Informational structure and contracts

A contract defines the transfer and actions for the worker and the firm within a match for all future histories. Call $s_\tau = (x_\tau, z_\tau) \in \mathbb{S} = \mathbb{X} \times \mathbb{R}$ the state of the match τ periods in the future and call $s^\tau = (s_1 ... s_\tau) \in \mathbb{S}^\tau$ a given history of realizations between s_1 the state today and s_τ, the state in τ periods.

The history of productivity is common knowledge to the worker and the firm and fully contractible. However the worker's actions are private information and transitions to other firms or to unemployment are assumed to be not contractible. This rules out side payments as well as countering outside offers[4]. Here, the contract offered by the firm to the worker is then represented by:

$$\mathcal{C} := (\mathbf{w}, \sigma); \quad \text{with } \mathbf{w} := \{w_\tau(s^\tau)\}_{\tau=0}^{\infty}, \text{ and } \sigma := \{v_\tau(s^\tau), e_\tau(s^\tau)\}_{\tau=0}^{\infty}, \tag{3.1}$$

I explicitly separate the firm's choice from the worker's response. The firm chooses the wage w_τ paid at every history and the worker responds by choosing (v_τ, e_τ) the search and effort decision[5]. σ can be thought as the action suggested by the contract and I will focus on contracts where the recommendation is incentive compatible. The contract space is completely flexible in the way it responds to tenure and any productivity history. In particular it leaves the firm free to chose how the wage should respond to productivity shock, which is the central question of this paper.

3.3.2 Worker choice

An unemployed worker of type x chooses optimally which sub-market (x, v_0) she applies to. The only value she cares about is the value she will get, specifically v_0 and the tightness of the market $\theta(x, v_0)$. Higher v_0 sub-markets

[4]Lentz (2013) develops a model with optimal contracts and countering of outside offers, but without productivity shocks, and shows that firms continue to backload wages.

[5]Derivations will later require a randomization which means that the contract can specify simple probability over actions instead of actions themselves. This is left implicit at this point but will be clarified in the recursive formulation of the problem.

deliver higher values but have longer average waiting times. I can write the value $\mathcal{U}(x)$ of being unemployed as follows:

$$\mathcal{U}(x) = \sup_{v_0 \in \mathbb{R}} b(x) + \beta p(\theta(x, v_0))v_0 + \beta \left(1 - p(\theta(x, v_0))\right) \mathbb{E}_{x'|x}\mathcal{U}(x'). \quad \text{(W-BE)}$$

We follow by writing the problem of the employed worker and the firm as a recursive contract. As presented in Spear and Srivastava (1987) the state space is augmented with V, the promised utility to the worker. The recursive contract is characterized at each (x, z, V) by $\{\pi_i, w_i, e_i, v_{1i}, W_{ix'z'}\}_{i=1,2}$ where $\pi_i : \mathbb{S} \times \mathbb{V} \to [0, 1]$ is a randomization, $w_i : \mathbb{S} \times \mathbb{V} \to \mathbb{R}_+$ is the wage, $e_i : \mathbb{S} \times \mathbb{V} \to [0, \bar{e}]$ is effort choice, $v_i : \mathbb{S} \times \mathbb{V} \to [0, \bar{v}]$ is the search choice and $W_{ix'z'} : \mathbb{S} \times (\mathbb{X} \times \mathbb{R}) \to \mathbb{V}$ is the utility promised for each realization next period.

The worker optimally chooses the action (v, e), when promised next period expected utility $W = \mathbb{E}_{x'z'}W_{x'z'}$, she solves the following problem:

$$\sup_{v,e} u(w) - c(e) + \delta(e)\beta\mathbb{E}_{x'|x}\mathcal{U}(x') + (1-\delta(e))\beta\kappa p(\theta(x, v))v + \beta(1-\delta(e))(1-\kappa p(\theta(x, v)))W,$$

for which we define the associated worker policies $v^* : \mathbb{X} \times \mathbb{V} \to [0, \bar{v}]$ and $e^* : \mathbb{X} \times \mathbb{V} \to [0, \bar{e}]$. Because of the properties of $p(\cdot)$, $\theta(\cdot, \cdot)$ and $c(\cdot)$, those functions are uniquely defined. Note that those policies only depend on the promised utility for next period and not on the current (z_t, V) as stated in the following definition.

Definition 5. *We defined the composite transition probabilities $\tilde{p} : \mathbb{X} \times \mathbb{V} \to \mathbb{R}$ and the utility return to the worker $\tilde{r} : \mathbb{X} \times V \to \mathbb{R}$ as functions of the promised utility W (using short-hand $e^* = e^*(x, W)$ and $v^* = v^*(x, W)$):*

$$p(x, W) = \kappa(1 - \delta(e^*)) \left(1 - p(\theta(x, v_1^*))\right) \quad (3.2)$$

$$\tilde{r}(x, W) = - c(e^*) + \beta\kappa \left(1 - \delta(e^*)\right) p(\theta(x, v_1^*)) \left(v_1^* - W\right) \quad (3.3)$$

$$+ \delta(e^*)\beta\mathbb{E}_{x'|x}U(x') + \beta \left(1 - \delta(e^*)\right) W. \quad (3.4)$$

(3.5)

These functions capture everything the firm needs to know about the consequences of setting the wage dynamically. We now turn to the firm's problem.

3.3.3 Firm profit, optimal contracting problem

I can now describe the firm problem in terms of promised utilities. The firm chooses a lottery over promised values and wages which then determines the participation probabilities. The expected profit of a match to the firm can be expressed recursively as

$$\begin{aligned}
\mathcal{J}(x,z,V) = & \sup_{\pi_i, w_i, W_i, W_{ix'z'}} \sum_{i=1,2} \pi_i \big(f(x,z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E}_{x'z'} \mathcal{J}(x', z', W_{ix'z'}) \big) \\
s.t \qquad & V = \sum_i \pi_i \left(u(w_i) + \tilde{r}(x, W_i) \right), \qquad \text{(BE-F)} \\
& W_i = \mathbb{E} W_{ix'z'}, \quad \sum \pi_i = 1.
\end{aligned}$$

The firm chooses the current period wage w_i and the promised utilities $W_{ix'z'}$ for each lottery realization i and state of (x', z') tomorrow. These control variables must be chosen to maximize expected returns subject to the promise keeping constraint. This constraint makes sure that the choices of the firm honors the promise made in previous periods to deliver the value V to the worker. The right hand side of the constraint is the lifetime utility of the worker given the choices made by the firm. The lottery is present only to insure concavity of the function.

The incentive compatibility of the worker is embedded in the $\tilde{r}$ and $\tilde{p}$ functions that we defined previously. By increasing future promises the firm can increase the probability that the match continues. However at given V, larger promised utilities go together with lower current wage w. Since the utility function is concave, there will be a point where too low of a wage is just not efficient. This is the classic insurance incentive tradeoff.

Finally firms choose how many vacancies to open in each (x, v) market. Given vacancy creation cost η and the fact that the match quality z starts at

0, the return to opening a vacancy is given by:

$$\Pi_0(x, V) = q(\theta(x, V))\mathcal{J}(x, 0, V) - \eta,$$

and firms will open vacancies in a given market if and only if expected profit is positive. The vacancy creation cost is linear, which means that if Π_0 is positive the firm will create an infinity of vacancies, if it's negative it won't create any and if it's zero the firm is indifferent.

3.3.4 Equilibrium definition

Free entry condition

We now impose a free entry condition on the market. Firms will open vacancies in each markets until the the expected profit is zero or negative:

$$\forall(x, V) \in \mathbb{X} \times \mathbb{V} : \quad \Pi_0(x, V) \leq 0. \tag{EQ1}$$

This will pin down the tightness of each market. $\phi(x, v)$ will denote the total mass of vacancies posted in market $(x.v)$.

Market clearing

Markets for labour must clear, in the sense that the equilibrium distribution must be generated by the equilibrium decisions. Given an equilibrium stationary distribution $h(x, y, z, V)$ of workers assigned to matches with a given promised utility, given the mass $\phi(x, V)$ of vacancies, the following clearing condition must be satisfied:

$$\begin{aligned}\forall x, v \quad \phi(x, v) = \theta(x, v)\Big[&u(x)\mathbf{1}[v_0^*(x) = v] \\ &+ \sum_{x \in \mathbb{X}} \int_z \int_{V'} \sum_i \pi_i(x, z, V')\mathbf{1}[v_i^*(x, W_i) = v] \; \mathrm{d}H(x, z, V')\Big].\end{aligned} \tag{EQ2}$$

There is one last market clearing equation for ϕ and it states that ϕ in the next period is consistent with itself, all the equilibrium decisions, and law motions such as the shocks on x, z and the endogenous separation. This is left for the appendix.

Definition 6. *A **stationary competitive search equilibrium** is defined by a mass of vacancies $\phi(x, v)$ across sub-markets (x, v), a tightness $\theta(x, v) \in \mathbb{R}$, an active job distribution $h(x, z, V)$ and an optimal contract policy $\xi = \{\pi_i, w_i, e_i, v_j, W_{ix'z'}\}_{i=1,2}$ such that:*

(a) ξ solves the firm optimal contract problem BE-F and so satisfies worker incentive compatibility.

(b) $\theta(x, v)$ and $\phi(x, v)$ satisfy the free entry condition EQ1 for all (x, v)

(c) $\theta(x, v)$, $\phi(x, v)$ and $h(x, z, V)$ solve the market clearing condition EQ2

(d) $h(x, z, V)$ is generated by $\phi(x, v)$ and ξ

The equilibrium assigns workers to firms with contracts in a way where neither workers or firms have an incentive to deviate. The distributions ϕ and h represent the equilibrium allocation.

3.3.5 Equilibrium and contract characterization

Lemma 17 (existence)**.** *A stationary competitive search equilibrium exists.*

Proof. See appendix 3.B.1 □

Menzio and Shi (2010) gives us the important results that a block recursive equilibrium exists in the version of this model with aggregate shocks and no worker effort or heterogeneity, and Tsuyuhara (2013) proves the existence with effort but without shocks or firm heterogeneity. The existence continues to be true when the incentive problem and the shocks are combined. The equilibrium is also well defined when adding aggregate shocks.

Lemma 18. *The Pareto frontier $\mathcal{J}(x, z, V)$ is continuously differentiable, decreasing and concave with respect to V and increasing in z.*

Proof. See appendix 3.B.3 □

Concavity is a direct implication of the use of the lottery. I then adapt the sufficient condition from Koeppl (2006) for differentiability in two-sided limited commitment models. From the free entry condition, the tightness function is a continuously differentiable and concave function of $\mathcal{J}(x, z, V)$, which implies that the composite search function $p(\theta(x, v))$ inherits those properties for all $x \in \mathbb{X}$.

I am interested in how firms decide to compensate workers over time given that they face the classic trade-off between insurance and incentives. The following proposition provides a clear prediction for how wages move dependent on the current state of the match:

Theorem 2 (optimal contract)**.** *For each viable match (x, z), independent of the lottery realization, the wage policy is characterized by a* ***target wage*** *$w^*(x, z)$, which is increasing in z such that:*

$$\begin{aligned} w_t \leq w^*(x_t, z_t) \quad &\Rightarrow \quad w_t \leq w_{t+1} \leq w^*(x_t, z_t) \quad \textit{incentive to search less} \\ w_t \geq w^*(x_t, z_t) \quad &\Rightarrow \quad w^*(x_t, z_t) \leq w_{t+1} \leq w_t \quad \textit{incentive to search more} \end{aligned}$$

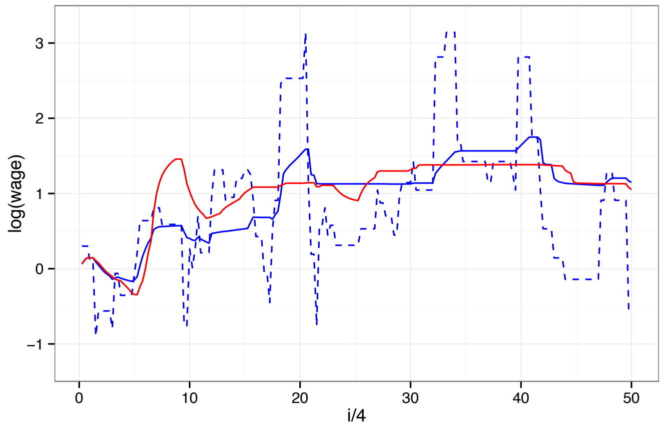
where the target wage is characterized by the zero expected profit condition for the firm:

$$\forall x, z \quad \mathbb{E}_{x'z'|xz}\mathcal{J}(x', z', W_{x'z'}) = 0$$

Proof. See Appendix 3.B.4. □

The optimal contract links wages to productivity. For all histories of shocks, the change in wage growth will be in the direction of the target wage which is itself tied to the productivity of the match. This means that workers' wages will respond to any shock affecting the expected productivity (Figure 3.2 shows an example wage path). In particular it will respond to both worker specific and firm productivity shocks. The exact change in the wage is char-

Figure 3.2: Wage and target wage example



Notes: This figure represents the target wage (dotted blue) and the actual wage (plain blue) for a worker. The red line represents a second worker sharing same firm specific shocks, but a different worker specific productivity.

acterized by the first order conditions of the firm problem (BE-F) and reads:

$$\forall x, z, x', z' \quad \frac{\tilde{p}_v(x, W_i)}{\tilde{p}(x, W_i)} \cdot \mathbb{E}_{z''y''}\mathcal{J}(x'', z'', W_{ix''z''}) = \frac{1}{u'(w_{x'z'})} - \frac{1}{u'(w)}.$$

The right hand side represents the change in marginal utilities and tells us that risk aversion affects how rapidly wages adjust. On the left hand-side the first term represents the severity of the moral-hazard problem and the second term is the discounted expected profit of the firm. This expression resembles the main equation in Rogerson (1985) and captures the incentive problem the firm is facing when paying the worker. When in a match the worker and the firm are part of a locally monopolistic bilateral relationship as in the original paper. However, the incentive problem here is precisely on the availability of the outside option. In Rogerson (1985), workers effort affects the output of the match whereas here, the effort affects its duration and the availability of outside options.

The fact that wages adjust downward even though firms can commit is the consequence of the existence of rents and the presence of an incentive problem.

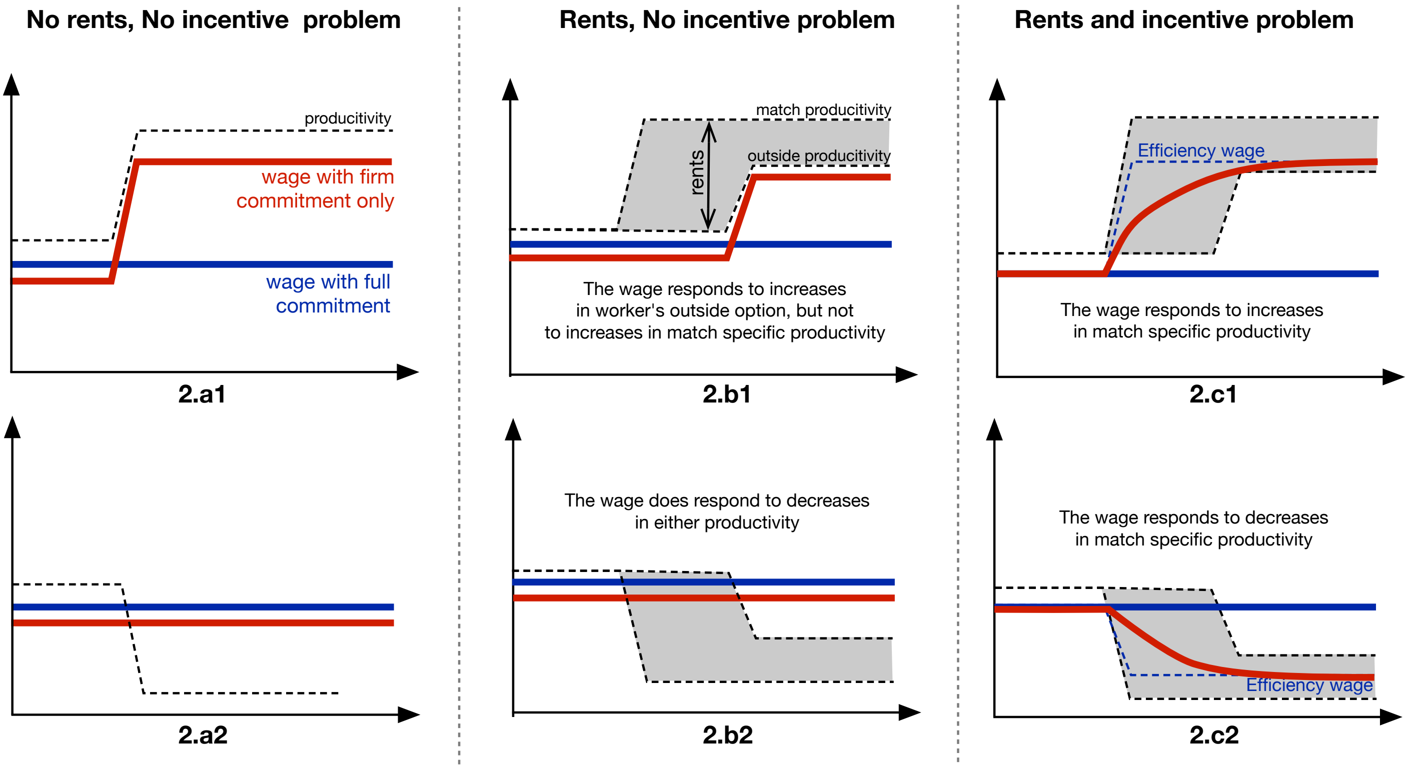


Figure 3.3: productivity and wages with rents, lack of commitment and incentive problems

In a competitive market without rents and with full commitment, even in the presence of productivity uncertainty, the firm will fully insure workers and the wage will be constant until the relationship is exogenously destroyed. The wage paid to the worker is the certainty equivalent of the present value of the firm output.

Harris and Holmstrom (1982) show that when allowing for only one-sided limited commitment, the wage will have to adjust when worker's productivity increases so as to retain her (Figure 3.3.a1). However negative shocks continue to be fully insured (Figure 3.3.a2). It is the lack of commitment that prevents the firm from offering the worker full insurance. Workers would want to commit ex-ante but can't and so the lack of commitment is a constraint, not a relaxation.

In the presence of rents the outside option of the worker and the productivity in the current match might vary separately. The outside option is linked to characteristics the worker caries with her when she moves to another firm, while the match rent also depends on the firm specific characteristics. Retaining the worker only requires offering more than her outside offer and and so only depends on worker-specific characteristics. This means that with rents only, the worker's wage does not respond to firm specific productivity shocks (Figure 3.3.b). Thomas and Worrall (1988) take the outside offers and the match rents as exogenous and add firm-side lack of commitment and show that in that case downwards adjustment will happen when the firm participation constraint binds. However in the interior region of the surplus, the contract fully insures the worker and the wage is constant.

The final ingredient is the incentive problem (Figure 3.3.c) which implies an unique efficient transfer from the firm to the worker instead of a full set. The worker chooses where to search and applies to increasingly long queues when promised higher values. Whenever the worker is getting less than the total value of the match, she will tend to leave the current job with a higher than efficient probability (inversely when the workers gets more, she does not search enough). Ex-ante, it is more efficient to sign a contract that will give up some of the insurance to come closer to the efficient worker decision. This dynamic was fully described in the extreme case of bilateral monopoly of Rogerson (1985) and continues to apply here in an equilibrium with firm competition and rents.

The continuum of queues available to the worker in the directed search equilibrium can be thought of as a probabilistic version of the constraint faced by firms in Harris and Holmstrom (1982). In their competitive version workers can find their $\bar{v}$ with probability one, whereas with directed search they can access any $v \leq \bar{v}$ with decreasing probability $p(v)$. In the presence of search frictions the firm-worker relationship becomes a temporary bilateral monopoly with an incentive problem determined by the equilibrium. As the strength of

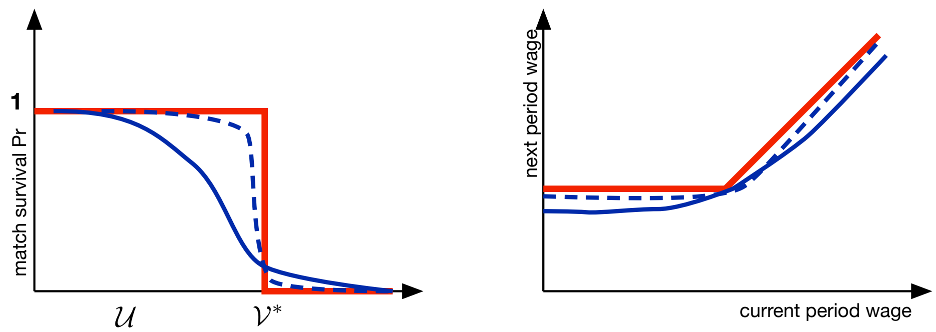


Figure 3.4: Meeting probability

the friction varies we get a continuum of contracts à la Rogerson (1985), with the property that as search frictions vanish, the contract becomes Harris and Holmstrom (1982) (See Figure 3.4).

Since rents and incentives are sufficient for the transfer of firm shocks to wages, search frictions are only one of several possible mechanisms. In the present model, there are two sources of rents, search frictions and match specific TFP, and two incentive problems, on the job search and effort choice e. This means that even when frictions are completely shut down, we would still see some firm level shocks in the earning dynamics[6]. Search frictions are an interesting feature not only because they allow us to consider employment risk but also because they generate both the rents and the incentive problem at the same time. It is also interesting to note that the shape of the meeting probability function creates some downward rigidities as in Harris and Holmstrom (1982).

Finally, firing never happens right away. First the firm decreases the wage of the worker over time because forcing her to search elsewhere is the most effective way for the firm to deliver ex-ante utility. The firm, when attracting the workers, can commit to paths where they keep the worker on payroll for

[6] I am in the process of estimating a frictionless version of the model on a subset of the moments that should be included in future version.

a given amount of time even though it means negative expected profit.

3.4 Estimation

3.4.1 Model specification and identification

I estimate the model using indirect inference and a parametrized model. I present in Table 3.4 the specification I use in the next sections. I use the constant relative risk aversion utility function. The discount rate for the worker and the interest rate for the firm are set to an annual 5% and the model is solved quarterly. The production function is parametrized by γ_a a scale parameter, γ_z and γ_x that control the dispersions in ability and match productivity. The worker effort function is such that $c(0) = 0, c'(\cdot) > 0, c''(\cdot) > 0$ and $\lim_{e \to 1} c(e) = \infty$. For the time being I set the flow value of unemployment to 30 percent of the starting productivity and I fix $c_1 = 0.3$ and $\gamma_z = 1$. I normalize the mean wage in the economy which pins down the value of γ_a. I also set an absolute lower bound of $-f(\bar{x}, z_0)/(10 \cdot r)$ on the negative surplus that firms can commit to. This leaves 6 parameters to estimate as shown in Table 3.6.

The vacancy cost η affects the meeting rate through the free entry condition (EQ1) and κ affects the relative efficiency of on-the-job search. The probability of exiting unemployment and the probability of job-to-job transitions pin down η and κ.

The effort cost function $c(\cdot)$ affects both the average rate at which workers loose their jobs and how this rates is linked to their current wage. c_0 and c_1 can be measured by fitting the slope and intercept of a logistic regression on the probability of employment to unemployment (E2U) transition conditional on current wage.

The parameter γ_x of the production function affects the return to worker ability x. Normalizing x to be uniform on $[0, 1]$ (at discrete uniformly spaced support), the production function f can be interpreted as the quantile function

matching function	$p(\theta) = \theta(1-\theta^{\nu})^{-1/\nu}$
utility function	$u(w) = \frac{w^{1-\varsigma}}{1-\varsigma}$
production function	$f(x,z) = \gamma_a \cdot \exp\left(\gamma_z \Phi^{-1}(z) + \gamma_x \cdot \Phi^{-1}(x)\right)$
worker cost function	$c(e) = c_0\left((1-e)^{-c_1} - 1\right)$
	$\delta(e) = 1 - e$
unemployment benefits	$b(x) = f(x, Q_z(b))$
worker type	$\Gamma_x(x_{t+1}\|x_t)$ is a Gaussian copula with parameter ρ_x
match TFP	$\Gamma_z(z_{t+1}\|z_t)$ is a Gaussian copula with parameter ρ_z
	updates to z_t are computed via ι_t shared at firm level

Table 3.4: functional form specifications

of worker specific heterogeneity. Using the normal distribution Φ^{-1} gives the simple interpretation that workers' productivity is distributed as a log-normal distribution with log-variance γ_x. The mean of that distribution is defined by γ_a which, as mentioned before, is normalized to match the mean log-wage in the economy.

The parameter of risk aversion controls how quickly changes in productivity get transmitted into wage changes. Every else kept equal, matching the total value added growth variance and the total wage growth variance within the firm gives an indication of how risk averse workers are.

Finally let's consider the parameters of the worker and match productivity processes. The values of ρ_x and ρ_z are learned from the variance of wage growth and the auto-covariance of wage growth among co-workers. The statistical model presented in the first section of the paper illustrates how the growth variance of worker is composed of both the worker specific growth and the firm specific growth and that the auto-covariance between co-workers' wage growth is mostly due to the common firm specific innovation. Matching both workers' wage growth variance and co-variance between co-workers allows to pin down ρ_x and ρ_z.

3.4.2 Solving the model

The model is estimated by method of simulated moments. For each parameter value I solve for the equilibrium, which is then used to simulate a representative sample. I create the moments from the simulated data and compute the weighted distance between the simulated moments and the moments measured from the Swedish data.

This approach requires resolving the model for each parameter set. I use a nested fixed point method where I jointly solve for the worker's problem, the firm's problem and the equilibrium constraint. The main difficulty resides in solving the firm problem where tackling directly (BE-F) requires finding the promised utilities $W_{z'x'}$ in each state of the world for the next period. This becomes infeasible as soon as reasonable supports are considered for $\mathbb{X}$ and $\mathbb{Z}$. However, the first order condition with respect to W reveals that the utility promised in different states are linked to each other. Call $\lambda\beta p(x, W)$ the multiplier for the $W = \sum W_{z'x'}$ constraint, then the first order condition for $W_{x'z'}$ is

$$\frac{\partial \mathcal{J}}{\partial V}(x', z', W_{x',z'}) = \lambda,$$

where given λ, if $\mathcal{J}$ is strictly concave, then all the $W_{x'z'}$ are pinned down. This reduces the search to one dimension. The simplification comes from the fact that the firm always tries to insure the worker as much as possible across future states, and does this by keeping her marginal utility constant across realizations. Indeed, we know that the derivative of $\mathcal{J}$ is the inverse marginal utility. One difficulty however is that $\mathcal{J}$ might be weakly concave in some regions. In that case one needs to keep track of a set of possible feasible promised utilities $W_{x'z'}$. Given the concavity of $\mathcal{J}$ this set will be an interval fully captured by its two extremities. This means that at worst the number of the control variables is augmented by one.

Using the marginal utility in the state space is known as the recursive Lagrangian approach as developed by Kocherlakota (1996); Marcet and Marimon (2011); Messner, Pavoni, and Sleet (2011); Cole and Kubler (2012). The problem of non-strict concavity persists in this formulation but Cole and Kubler (2012) show how to overcome this difficulty by keeping track of the upper and lower bound of the set of solutions. Numerically I solve the firm problem using recursive Lagrangian and do not find any such flat region. The recursive Lagrangian for the firm problem is derived in Appendix 3.B.6 and is given by:

$$\begin{aligned}\mathcal{P}(x,z,\rho) = \inf_{\gamma} \sup_{w,W} f(x,z) - w + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\ - \beta\gamma\tilde{p}(x,W) + \beta\tilde{p}(x,W)\mathbb{E}\mathcal{P}(x',z',\gamma), \quad (3.6)\end{aligned}$$

where

$$\mathcal{P}(x,z,\rho) := \sup_{v} \mathcal{J}(x,z,v) + \rho v.$$

3.4.3 Estimation and standard errors (Preliminary)

Estimation of the parameters is achieved using a minimum distance estimator based on a set of moments m_n. The method is close to simulated moments, however because of the moments are based on individual data and some are based on aggregation at the firm level, I present it as an indirect inference estimator.

Definition 7. *Given a vector m_n of moments such that $\sqrt{n}\left(m_n - m(\theta_0)\right) \xrightarrow{d} \mathcal{N}(0,\Sigma)$ where θ_0 is the true parameter, and for a given weighting matrix $W_n = O(1)$, I define the following criterion:*

$$L_n(\theta) = -\frac{n}{2}\left[m_n - m(\theta)\right]^T W_n \left[m_n - m(\theta)\right],$$

and the associated minimum distance estimate $\hat{\theta}_n = \inf_\theta L_n(\theta)$.

Because some of the moments are defined at the firm level, such as the correlation between co-worker wage growth, n refers to the number of firms. Point estimates are computed using a parallel version of differential evolution, see Das and Suganthan (2011) for a complete survey. In the first stage I use a weighting matrix constructed from the inverse diagonal of an estimate from the data of Σ which ignores the serial correlation and the fact that the same worker appears in several firms:

$$W_n = \left(\text{diag}\left[\hat{\Sigma}\right]\right)^{-1}.$$

The computation of standard errors is based on the pseudo-likelihood estimator presented in Chernozhukov and Hong (2003). Using MCMC rejection sampling, I can perform the estimation in parallel, without having to compute derivatives and still obtain standard errors on the parameters. Given the criterion $L_n(\theta)$, with moments m_n, true parameter θ_0 and weighting matrix W_n, the asymptotic variance for the minimum distance estimator $\hat{\theta_n}$ is distributed according to

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{d} \mathcal{N}(\theta_0, J^{-1}\Omega J^{-1})$$

where

$$\begin{aligned} \Omega &= \lim_{n\to\infty} \quad \left[\frac{\partial m(\theta_0)}{\partial \theta^T}\right]^T W_n \Sigma W_n \frac{\partial m(\theta_0)}{\partial \theta^T} \\ J &= \lim_{n\to\infty} \quad \frac{1}{n}\frac{\partial^2 L_n(\theta)}{\partial \theta^T \partial \theta} \end{aligned}$$

The full procedure requires two steps. In a first step I acquire a consistent estimate of $\hat{\theta}_n$ using an approximate weighting matrix $\hat{\Sigma}_n$ using bootstrap. Given a good value of $\hat{\theta}_n$ I compute a Markov chain from the posterior of the pseudo likelihood of $L_n(\theta)$ as described Chernozhukov and Hong (2003) and extended to parallel chains as presented in Baragatti, Grimaud, and Pommeret (2011). The Markov chain allows construction of an estimate of Ω and J^{-1}. J^{-1} is obtained by taking the variance covariance matrix of the parameters generated by the chain. Ω can be computed by finite differences around the optimal value $\hat{\theta}_n$ by selecting draws from the chain that are close to it. A consistent estimate of Σ can then be constructed by simulating the model at

	HS dropout		HS grad		Some college	
	model	data	model	data	model	data
Pr_{U2E}	0.131	0.152 (2.69e-04)	0.214	0.184 (1.53e-04)	0.209	0.191 (3.36e-04)
Pr_{J2J}	0.0224	0.0223 (4.01e-05)	0.0284	0.0267 (2.27e-05)	0.0338	0.0331 (3.21e-05)
Pr_{E2U}	0.0202	0.0249 (6.25e-05)	0.0199	0.0223 (2.67e-05)	0.0164	0.0143 (3.05e-05)
$E(\Delta \log w_{it} \| EE)$	0.0145	0.0125 (1.73e-04)	0.03	0.0153 (8.69e-05)	0.0257	0.0335 (1.24e-04)
$E(\Delta \log w_{it} \| J2J)$	0.0329	0.0274 (8.36e-04)	0.0738	0.0306 (3.95e-04)	0.0875	0.0506 (5.60e-04)
$Var(\log w_{it})$	0.163	0.127 (2.09e-04)	0.141	0.116 (1.32e-04)	0.204	0.203 (3.38e-04)
$Var(\Delta \log w_{it} \| EE)$	0.0186	0.0198 (2.38e-05)	0.0171	0.0173 (1.61e-05)	0.0173	0.0193 (2.42e-05)
$Var(\Delta \log w_{it} \| J2J)$	0.0448	0.0206 (5.36e-04)	0.0353	0.018 (2.14e-04)	0.0466	0.0186 (2.49e-04)
$Var(\Delta \log y_{it})$	0.375	0.103 (1.24e-03)	0.158	0.119 (1.10e-03)	0.102	0.132 (1.65e-03)
$Cov(\Delta \log w_{it}, \Delta \log w_{jt} \| EE)$	0.00154	0.00126 (2.64e-06)	0.00169	0.00167 (1.80e-06)	0.0023	0.00235 (3.32e-06)

Table 3.5: Within sample model fit (Preliminary)

$\hat{\theta}_n$ and computing the covariance matrix.

3.4.4 Moments and estimates

I present here the set of moments used for estimation on the different education groups. Table 3.5 reports the moments in the data with their measured standard deviation and the value of the moments in the model at the estimated parameter values. Table 3.6 presents the estimated parameters for each education group.

The model matches transition probabilities and variances quite precisely across education groups. However at this time the model performs poorly on the average wage growth on the job and the mean wage gain on job-to-job transitions. Those moments are related to each other because the job-to-job transition rate, mean gain on moving and on the job mean wage growth are linked to each other because wages increase on-the-job to lower the worker search decision. This is a common limitation of search model which suggests

		HS dropout	HS grad	Some college
scale (log wage)		0.127 (0.000209)	10.4 (0.0003)	10.4 (0.0003)
risk aversion	ς	1.12 (0.124)	1.62 (0.0408)	1.42 (0.0586)
vacancy cost	η^-1	1.34 (0.34)	0.646 (0.0753)	0.605 (0.0532)
OTJ efficiency	κ	0.586 (0.15)	0.617 (0.0238)	0.687 (0.0387)
effort cost	c_0	0.0779 (0.0244)	0.0498 (0.0202)	0.0418 (0.0229)
worker heterogeneity	γ_x	2.03 (0.303)	1.27 (0.123)	1.5 (0.0797)
worker type auto-cor	ρ_x	0.749 (0.0365)	0.802 (0.0206)	0.879 (0.0274)
match type auto-cor	ρ_z	0.765 (0.06)	0.962 (0.0502)	0.978 (0.0215)

Table 3.6: Parameter estimates (Preliminary)

that some human capital accumulation might be happening in the data. This is absent from the current model.

3.5 Empirical implications

3.5.1 Decomposition of permanent wage growth

I can now utilize the model to decompose observed variances into better defined welfare measures. Our concern is with the sources of uncertainty in the change of lifetime utility, however to get measures in monetary form, I define the wage growth variance of log permanent wage as:

$$\mathbb{E}_t \left(\bar{w}_{t+1} - \bar{w}_t\right)^2 \quad \text{where} \quad \bar{w}_t := \log\left(u^{-1}\left(rW_t\right)\right),$$

where $\bar{w}$ represents the annuity wage that delivers the current level of lifetime utility, the permanent wage equivalent to the expected lifetime utility. This is a meaningful measure since W_t includes all possible future risk of loosing the job or the opportunities to find new ones. Similarly we can measure equivalent permanent output that I will denote $\bar{y}$. Considering employed workers, five mutually exclusive events can happen to them over the course of a period: i) job loss, ii) job transition, iii) firm shock, iv) worker shock or v) none of the above. We can decompose the permanent earning growth variance into the contributions of those five events:

$$\mathbb{E}_t \left(\bar{w}_{t+1} - \bar{w}_t\right)^2 = \sum_{i=1}^{5} p(\mathrm{ev}_i) \cdot \mathbb{E}_t \left[\left(\bar{w}_{t+1} - \bar{w}_t\right)^2 |\mathrm{ev}_i\right] = \sum_{i=1}^{5} V_i.$$

To get the average risk in the population, I integrate the V_i over the stationary distribution. Table 3.7 reports this variance decomposition for the three education groups. Including $p(\mathrm{ev}_i)$ in the computation of V_i directly accounts for the likelihood of the event.

To get an idea of the overall underlying uncertainty I compute a pass through measure that links the growth variance in productivity to the growth variance in earnings:

$$\frac{Cov(\bar{w}_{t+1} - \bar{w}_t, \bar{y}_{t+1} - \bar{y}_t)}{Var(\bar{y}_{t+1} - \bar{y}_t)}$$

and report this value conditional on receiving a worker shock and firm shock and unconditional.

The results first tell us that the total uncertainty associated with mobility is of the same magnitude as the uncertainty associated with productivity shocks. For high school drop out mobility accounts for 50 percent of uncertainty and for 24 percent for college graduates. Within mobility, job loss takes a bigger share for high school drop outs than for college graduates. This seems intuitive given the $J2J$ and $E2U$ transition rates of the two groups. Among job stayers, firm productivity shocks represent the main source of uncertainty.

Finally the pass through measure indicates that even though different education group suffer differently from firm and worker shock in terms of total earning uncertainty, the way in which those uncertainty transmit seems to be the same. For both education groups, a 10 percent change productivity due to a firm shock generates a 3 percent drop in permanent earnings. Similarly a 10 percent drop in productivity due to a worker shocks translates into a 2

	HS dropout		HS grad		Some college	
	Growth variance shares					
firm shock	4.6e-04	19.6%	1.5e-04	17.8%	3.1e-04	19.2%
worker shock	1.3e-03	54.2%	2.9e-04	34.5%	7.1e-04	44.4%
job change	1.7e-04	7.13%	1.2e-04	13.8%	1.8e-04	11.3%
job loss	4.2e-04	18.1%	2.8e-04	32.4%	3.9e-04	24.3%
no shock	2.3e-05	0.968%	1.2e-05	1.36%	1.5e-05	0.933%
	Passthrough coefficents					
overall	0.369		0.243		0.282	
worker shock	0.388		0.179		0.215	
firm shock	0.348		0.271		0.328	

Table 3.7: Permanent wage growth variance decomposition

percent drop on average.

3.5.2 Policy analysis

I analyze the effect of a revenue neutral government policy that redistributes from high wages to lower wages. I parametrize the policy as follows:

$$\tilde{w} = \lambda w^{\frac{1}{\tau}}.$$

I use the highest education group for the analysis, fix $\lambda = 1.2$ and solve for $\tau = 1.25$ to make the policy revenue neutral. To get a better understanding of the effect of the policy, I report four sets of numbers: i) the model solved at the estimated parameters, without any transfer, ii) use the same solution and apply transfers without adjusting decisions, iii) solve the model again with agents knowing about the transfers, and report pre-transfer moments and iv) post-tax moments. Figure 3.5 represents graphically the transfer and Table 3.8 reports the computed results.

The goal of the policy is to reduce both the uncertainty in earnings growth and the cross-sectional inequality. When applied directly on the equilibrium solution we see that total log wage variance is reduced by 36%, and the wage growth variance is reduced by 35%. However agents react to the introduction

Agents	do not expect transfers		expect transfers	
Transfers	before	after	before	after
Output	1		0.978	
Unemployment	4.96%		4.69%	
Total wage	1	0.992	0.962	0.961
Wage variance	0.205	0.131	0.238	0.152
Growth variance	0.0174	0.0111	0.021	0.0134

Table 3.8: Revenue neutral policy

of the policy in a way that attenuates its direct effect. Re-solving the model including those transfers gives a reduction in log wage variance of only 10% and for wage growth of 30%.

The policy however also affects unemployment which goes from 4.96% to 4.56%. This happens because the policy makes lower productivity jobs marginally more productive than without transfers, favoring workers coming out of unemployment who apply to lower paying, highly accessible jobs. On the other hand total output is reduced for a similar reason, worker reallocation is not as critical and in the economy with transfers, worker will reallocate less efficiently.

3.6 Conclusion

In this paper I study the different sources of uncertainty faced by workers in the labour market. Workers are subject to individual productivity shocks and their earnings may also be affected by the performance of their employer because of search frictions in the labour market. To understand the way shocks get transmitted and how this might affect welfare and labour market policy I develop an equilibrium model with search frictions, risk averse workers, firm and worker productivity shocks. In this model I show that the optimal contract pays a wage that smoothly tracks the joint match productivity. This implies that both worker and firm level shocks transmit to wages, albeit only partially. In contrast to the perfectly competitive model, on one hand firm may insure workers' productivity shocks but on the other hand they are able to transmit firm level shocks to wages.

I estimate the model on matched employer-employee data to estimate the relative importance of different sources of uncertainty. Firm productivity shocks can account for 20% for the overall permanent wage uncertainty, leaving mobility and worker shocks as the main sources of risk. Firms are unable to insure workers once the employment relationship ends making publicly provided unemployment benefits an important source of insurance. To quantify the underlying source of uncertainty I compute a pass-trough measure of productivity shocks to earning shocks and find that 20% of worker shocks and 30% of firm shocks get transmitted to wages. The implication of those findings is that policies should focus on transitions in and out of work. This is because when employed the firm will provide some source of insurance, but the firm can't continue to insure the worker when the relationship ends.

An important extension to this model is to allow individuals to hold assets, which would allow them to self insure. The inclusion of observable assets would depart only slightly from the current version of the model but a more realistic environment would allow workers to privately save. This creates many interesting economic questions such as how do firms recruit among workers with different asset holdings? In preliminary analysis of such an extension I find that firms try to hire workers with higher assets because they are easier to incentivize: firms can backload even more or get them to pay a bond, improving retention. Upfront payment by the worker to the firm is observed in high skill labour markets such as partnerships in law and consulting firms.

Another extension is to allow firms to counter outside offers. Inefficient poaching happens rarely in the estimated version of the model, but it would be more realistic to have a mechanism by which firms could optimally decide whether to counter outside offers. This type of negotiations happen in practice in high skills markets such as CEO and academics.

Bibliography

Altonji, J., A. Smith, and I. Vidangos (2009): "Modeling earnings dynamics," *NBER*.

Attanasio, O. P., and N. Pavoni (2011): "Risk sharing in private information models with asset accumulation: Explaining the excess smoothness of consumption," *Econometrica*, 79(4), 1027–1068.

Azariadis, C. (1975): "Implicit contracts and underemployment equilibria," *The Journal of Political Economy*, pp. 1183–1202.

Baily, M. N. (1974): "Wages and employment under uncertain demand," *The Review of Economic Studies*, 41(1), 37–50.

Baragatti, M., A. Grimaud, and D. Pommeret (2011): "Parallel Tempering with Equi-Energy Moves," *Arxiv preprint*.

Benveniste, L., and J. Scheinkman (1979): "On the differentiability of the value function in dynamic models of economics," *Econometrica: Journal of the Econometric Society*, pp. 727–732.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *The American Economic Review*, pp. 1887–1921.

Burdett, K., and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71(5), 1377–1404.

Burdett, K., S. Shi, and R. Wright (2001): "Pricing and matching with frictions," *Journal of Political Economy*, 109(5), 1060–1085.

Chernozhukov, V., and H. Hong (2003): "An MCMC approach to classical estimation," *Journal of Econometrics*, 115(2), 293–346.

Cole, H., and F. Kubler (2012): "Recursive contracts, lotteries and weakly concave pareto sets," *Review of Economic Dynamics*.

Dardanoni, V. (1995): "Income distribution dynamics: monotone Markov chains make light work," *Social Choice and Welfare*, 12(2), 181–192.

Das, S., and P. N. Suganthan (2011): "Differential evolution: A survey of the state-of-the-art," *Evolutionary Computation, IEEE Transactions on*, 15(1), 4–31.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the Firm.," *Journal of Political Economy*, 113(5), 34.

Harris, M., and B. Holmstrom (1982): "A theory of wage dynamics," *The Review of Economic Studies*, 49(3), 315–333.

Hopenhayn, H., and J. Nicolini (1997): "Optimal unemployment insurance," *Journal of political economy*, 105(2), 412–438.

Knight, F. H. (1921): "Risk, uncertainty and profit," *New York: Hart, Schaffner and Marx*.

Kocherlakota, N. (1996): "Implications of efficient risk sharing without commitment," *The Review of Economic Studies*, 63(4), 595.

Koeppl, T. (2006): "Differentiability of the efficient frontier when commitment to risk sharing is limited," *Topics in Macroeconomics*, 6(1), 1–6.

Lentz, R. (2013): "Optimal Wage-Tenure Contracts without Search Intensity Commitment," *Working Paper*.

Low, H., C. Meghir, and L. Pistaferri (2009): "Wage risk and employment risk over the life cycle," *NBER*.

MaCurdy, T. E. (1982): "The use of time series processes to model the error structure of earnings in a longitudinal data analysis," *Journal of econometrics*, 18(1), 83–114.

Marcet, A., and R. Marimon (2011): "Recursive contracts," *Unpublished Manuscript*.

MENZIO, G., AND S. SHI (2009): "Efficient search on the job and the business cycle," *Unpublished Manuscript*.

——— (2010): "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 145(4), 1453–1494.

MESSNER, M., N. PAVONI, AND C. SLEET (2011): "Recursive methods for incentive problems," *Working Papers*.

MOEN, E. (1997): "Competitive search equilibrium," *Journal of Political Economy*, pp. 385–411.

MONTGOMERY, J. (1991): "Equilibrium wage dispersion and interindustry wage differentials," *The Quarterly Journal of Economics*, 106(1), 163–179.

PETERS, M. (1991): "Ex ante price offers in matching games non-steady states," *Econometrica: Journal of the Econometric Society*, pp. 1425–1454.

ROGERSON, W. P. (1985): "Repeated moral hazard," *Econometrica*, pp. 69–76.

ROYS, N. (2011): "Persistence of Shocks, Employment and Wages," *Unpublished Manuscript*.

RUDANKO, L. (2009): "Labor market dynamics under long-term wage contracting," *Journal of Monetary Economics*, 56(2), 170–183.

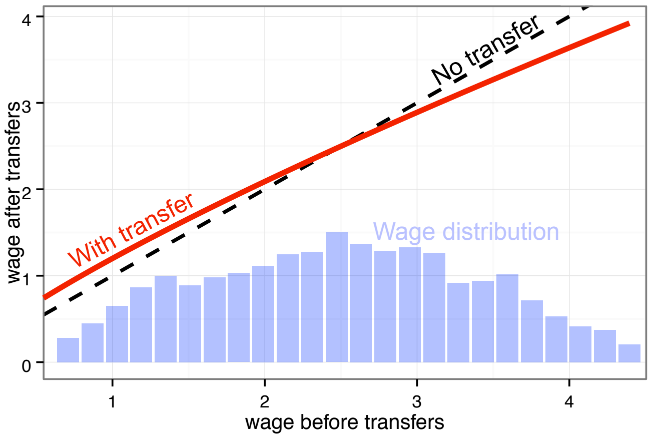
SCHAAL, E. (2010): "Uncertainty, Productivity and Unemployment during the Great Recession," *unpublished, December*.

SHI, S. (2009): "Directed Search for Equilibrium Wage–Tenure Contracts," *Econometrica*, 77(2), 561–584.

SHIMER, R. (1996): "Contracts in a frictional labor market," Discussion paper, mimeo.

SHIMER, R. (2001): "The assignment of workers to jobs in an economy with coordination frictions," *NBER*.

Figure 3.5: Policy



Notes: This figure represents the pre and post tranfer wages together with the distribution of wages for the polocy experiment considered here $\tilde{w} = \lambda w^{\frac{1}{\tau}}$.

SPEAR, S. E., AND S. SRIVASTAVA (1987): "On repeated moral hazard with discounting," *The Review of Economic Studies*, 54(4), 599–617.

THOMAS, J., AND T. WORRALL (1988): "Self-enforcing wage contracts," *The Review of Economic Studies*, 55(4), 541–554.

TSUYUHARA, K. (2013): "Dynamic Contracts with Worker Mobility via Directed On-the-Job Search," *Working Paper*.

VILLANI, C. (2003): *Topics in optimal transportation*, vol. 58. AMS Bookstore.

3.A Data

3.A.1 Auxiliary model

Recall the auxiliary model described in the first section of the paper. Note that δ_{jt} appears alone in the very first model, but is then decomposed into two different components when value added is introduced.

$$\begin{aligned}
w_{ijt} &= \beta Z_t + \tilde{w}_{ijt} + v_{ijt} \\
\tilde{w}_{ijt} &= \tilde{w}_{ijt-1} + \delta_{jt} + \xi_{ijt}, \\
y_{jt} &= \beta X_t + \tilde{y}_{jt} + u_{jt} \\
\tilde{y}_{jt} &= \tilde{y}_{jt-1} + \mu_{jt} \\
\delta_{jt} &= \tau \mu_{jt} + \nu_{jt}
\end{aligned}$$

The auxiliary model presented can be recovered from the following moments:

$$\mathbb{E}_j \left[(\mathbb{E}_i \Delta w_{ijt})^2 \right] = \sigma_\delta^2 = \sigma_\nu^2 + \tau^2 \sigma_\mu^2 \tag{m1}$$

$$\mathbb{E}_{ij} \left[(\Delta w_{iit})^2 \right] = \sigma_\xi^2 + \sigma_\delta^2 + 2\sigma_v^2 \tag{m2}$$

$$\mathbb{E}_{ij} \left[(\Delta y_{it})^2 \right] = \sigma_\mu^2 + 2\sigma_u^2 \tag{mw1}$$

$$\mathbb{E}_j \left[\Delta y_{jt} \cdot \Delta y_{jt-1} \right] = -\sigma_v^2 \tag{mw2}$$

$$\mathbb{E}_j \left[\Delta y_{ijt} \cdot \Delta w_{ijt} \right] = \tau \sigma_\mu^2 \tag{mw2}$$

where $\mathbb{E}_i$ represents the expectation over co-workers within firm j.

3.B Proofs

3.B.1 Existence of the equilibrium

The model presented here is similar to the one presented in Menzio and Shi (2010). The differences are the composite functions $\tilde{r}$ and $\tilde{p}$ that now include the effort decision and the fact that workers are now heterogenous. This means that I can apply their proof here as long as I can derive the necessary properties on $(\tilde{p}, \tilde{r})$ and show that heterogeneity does not break any of the Lipschitz bounds.

Lemma 17 (existence)**.** *A stationary competitive search equilibrium exists.*

Definition 8. *call $\mathbb{J}$ the set of functions $\mathcal{J} : \mathbb{X} \times \mathbb{Z} \times \mathbb{V} \to \mathbb{R}$ such that*

(a) $\mathcal{J}$ is strictly decreasing in V,

(b) bi-Lipschitz continuous in V

(c) bounded

(d) concave

Lemma 19. *The operator T defined in* (BE-F) *is self-mapping on $\mathbb{J}$.*

Proof of Lemma 17 . Consider a function $\mathcal{J} \in \mathbb{J}$ and its image $\hat{\mathcal{J}} = T\mathcal{J}$. We start by noting that the lottery gives us that $\hat{\mathcal{J}}$ is concave which gives continuity and almost everywhere differentiability. Given that, we can apply the envelope theorem to find that the derivative of $\hat{\mathcal{J}}$ is almost everywhere $-1/u'(w^*(x, y, V))$. Given that we have established that the offered wage has to be bounded, it gives that the derivative of $\hat{\mathcal{J}}$ is also bounded in $[-1/\bar{u}', -1/\underline{u}']$. Given that $\mathbb{V}$ is itself bounded it gives us that $\hat{\mathcal{J}}$ is also bounded. The derivative is also strictly negative and so $\hat{\mathcal{J}}$ is a one to one mapping. $\hat{\mathcal{J}}$ is then also bi-Lipschitz. That concludes the fact that $\hat{\mathcal{J}} \in \mathbb{J}$. $\square$

Lemma 20. *Bounds on $\tilde{p}$,$\tilde{r}$ [incomplete]*

First I report a result from Menzio and Shi (2010) which applies directly here and states that given $\mathcal{J}_n, \mathcal{J}_r$ such that $||\mathcal{J}_n - \mathcal{J}_r|| < \rho$ we have that $\forall x, v$

$$\begin{aligned}
\|p(\theta(x, v_{1n}^*)) - p(\theta(x, v_{1r}^*))\| &< \alpha_P(\rho) = \max\{2\bar{B}_P + p'(0)\alpha_\theta\rho, 2\alpha_R\rho^{1/2}\} \\
\|p(\theta(x, v_{1n}^*))\,(v_{1n}^* - v) - p(\theta(x, v_{1r}^*))\,(v_{1r}^* - v)\| &< \alpha_R\rho
\end{aligned}$$

that we need to use to show that it continues to apply when the effort choice of the worker is added. Given the policy for job search the effort choice is given by $\delta = e = c'^{-1}(p(\theta(x, v_{1n}^*))\,(v_{1n}^* - v) + v - \mathcal{U}(x))$ and so given that v itself is bounded we find new bounds on the $\tilde{p}$ and $\tilde{r}$ functions:

$$\begin{aligned}
\|\tilde{r}_n - \tilde{r}_r\| &< \alpha_r\rho \\
\|\tilde{p}_n - \tilde{p}_r\| &< \alpha_P(\rho)
\end{aligned}$$

Lemma 21. *The operator T is continuous on $\mathbb{J}$*

Proof. This boils down to showing that T is K-lipschitz. Let's take two functions $\mathcal{J}_1, \mathcal{J}_2 \in \mathbb{J}$ and their respective image $\hat{\mathcal{J}}_1, \hat{\mathcal{J}}_2$. We already know that they are part of $\mathbb{J}$. Then we need to find a constant K such that $||\hat{\mathcal{J}}_1 - \hat{\mathcal{J}}_2|| \leq K||\mathcal{J}_1 - \mathcal{J}_2||$. We substitute in the $\hat{\mathcal{J}}_1$ and $\hat{\mathcal{J}}_2$ by their definition. We then bound each element separately:

$$\begin{aligned}||\hat{\mathcal{J}}_1(x,z,V) - \hat{\mathcal{J}}_2(x,z,V)|| \leq ||u(w_1) - u(w_2)|| \\ + ||\tilde{p}_1(x,W_1)\mathbb{E}\mathcal{J}(x',z',W_{1x'z'}) - \tilde{p}_2(x,W_2)\mathbb{E}\mathcal{J}(x',z',W_{2x'z'})||\end{aligned}$$

where we now want to bound each term.

The rest of the proof follows identical steps to Tsuyuhara (2013) and Menzio and Shi (2010). □

3.B.2 Properties or worker search functions

Lemma 22. *Given (x, W), $v^*(x, W)$ and $e^*(W)$ are uniquely determined, $\tilde{p}(x, W)$ is continuous and decreasing, $\tilde{r}(x, W)$ is increasing in W, continuously differentiable and $\frac{\partial \tilde{r}}{\partial W}(x, W) = \beta\tilde{p}(x, W)$.*

Proof. remember the definitions

$$\begin{aligned}v^*(x,W) =& \arg\max_v p(\theta(x,v))(v - W) \\ e^*(x,W) =& \arg\max_e -c(e) + \delta(e)\beta\mathbb{E}W_0(x') \\ &+ \beta(1-\delta(e))\left(p(\theta(x,v^*))v^* + \beta(1-\delta(e))(1-p(\theta(x,v^*)))W\right),\end{aligned}$$

and the definition of the composite functions

$$\begin{aligned}\tilde{p}(x,W) =& (1-\delta(e^*(x,W)))\left(1-p(\theta(x,v_1^*(x,W)))\right) \\ \tilde{r}(x,W) =& -c(e^*(x,W)) + \beta(1-\delta(e^*(x,W)))p(\theta(x,v_1^*(x,W)))\left(v_1^*(x,W) - W\right) \\ &+ \delta(e^*(x,W))\beta\mathbb{E}_{x'|x}U(x') + \beta(1-\delta(e^*(x,W)))(x,W)W\end{aligned}$$

I first normalize $\delta(e) = 1 - e$ (or equivalently redefine c and e such that

$c(e) = c(\delta^{-1}(e)))$, where $c(e)$ is increasing and concave. The maximization problem for v gives the following first order condition

$$p'(\theta(x,v))(v-W) + p(\theta(x,v)) = 0$$

where given the property of p and q and the equilibrium definition of θ we have that the function $v \mapsto p(\theta(x,v))$ is decreasing and strictly concave. This gives that the maximum is unique and so $v^*(x,W)$ is uniquely defined. The first order condition for e is given by

$$c'(e) = \beta p(\theta(x, v_1^*(x,W)))\,(v_1^*(x,W) - W) + \beta W - \beta \mathbb{E}_{x'|x} U(x')$$

and given the assumption that c is strictly convex, we get that $e^*(x,W)$ is also uniquely defined.

Finally we can use the envelope condition to compute the derivative of $\tilde{r}$ with respect to W. By definition we have

$$\tilde{r}(x,W) = \sup_{v,e} u(w) - c(e) + (1-e)\beta \mathbb{E}_{x'|x} W_0(x') + e\beta p(\theta(x,v))v + e\beta(1-p(\theta(x,v)))W,$$

and so we get

$$\frac{\partial \tilde{r}}{\partial W}(x,W) = \beta e^*(x,W)(1-p(\theta(x,v^*(x,W))) = \beta \tilde{p}(x,W)$$

which proves that $\tilde{r}$ is continuously differentiable as long as $\tilde{p}$ is continuous. □

3.B.3 Regularity properties for equilibrium functions

Lemma 18. *The Pareto frontier $\mathcal{J}(x,z,V)$ is continuously differentiable, decreasing and concave with respect to V and increasing in z.*

Proof of Lemma 18 . Consider the optimal contract equation:

$$
\begin{aligned}
\mathcal{J}(x,z,V) = \sup_{\pi_i, W_i, W_{ix'y'}} & \sum \pi_i \big(f(x,z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E} \mathcal{J}(x', z', W_{ix'y'})\big) \\
s.t \quad (\lambda) \quad & 0 = \sum_i \pi_i \left(u(w_i) + \tilde{r}(x, W_i)\right) - V, \\
(\gamma_i) \quad & 0 = W_i - \mathbb{E} W_{ix'y'}, \\
& \sum \pi_i = 1.
\end{aligned}
$$

We already know that $\mathcal{J}$ is concave because of the two point lottery. That tells us that it is continuous and differentiable almost everywhere. Let's then show that it is differentiable everywhere. I follow the steps of the derivation presented in Koeppl (2006) where he shows that in the problem with two sided limited commitment it is sufficient to have one state realization where neither participation constraint binds to achieve differentiability of the Pareto frontier. Given that the current problem is one sided the result works almost right away, it just needs to be extended to include a search decision.

For a fixed $s = (x, z)$, let's consider a point $\tilde{V}$ where it's not differentiable and call $(\tilde{w}, \tilde{\pi}_1, \tilde{W}_{ix'z'}, \tilde{W}_i)$ the firm's action at that point. This action is by definition feasible and delivers $\tilde{v}$ to the worker. From that strategy I am going to construct a continuum that delivers any V around $\tilde{V}$. Keeping $(\tilde{\pi}_1, \tilde{W}_{ix'z'}, \tilde{W}_i)$ the same, I defined $w^*(V) = u^{-1}(V - \tilde{V})$.

I then define the function $\tilde{\mathcal{J}}(s, v)$ as the value that uses strategy $\left(w^*(V) = u^{-1}(V - \tilde{V}), \tilde{\pi}_1, \tilde{W}_{ix'y'}, \tilde{W}_i\right)$. It is the case that the strategy is feasible since all constraints remain satisfied. By definition of $\mathcal{J}$ we have that $\tilde{\mathcal{J}}(s, V) \leq \mathcal{J}(s, V)$ together with $\tilde{\mathcal{J}}(s, \tilde{V}) = J(s, \tilde{V})$. Finally because $u(\cdot)$ is concave, increasing and twice differentiable, $\tilde{\mathcal{J}}(s, \tilde{V})$ is also concave and twice differentiable.

We found a function concave, continuously differentiable, lower than $\mathcal{J}$ and equal to $\mathcal{J}$ at $\tilde{V}$ we can apply Lemma 1 from Benveniste and Scheinkman (1979) which gives us that $\mathcal{J}(s, v)$ is differentiable at $\tilde{v}$. We then conclude that

$\mathcal{J}$ is differentiable everywhere. Finally let's show that $\mathcal{J}(x,z,v)$ is increasing in z.

Let's consider two different values $z_1 < z_2$. Call ξ_i the history contingent policy starting at (x, z_i, v). Policy ξ_1 will deliver identical utility to the worker in all histories independently of whether it started at z_1 or z_2. I then compare the value of using ξ_1 at (x, z_1, v) and (x, z_2, v). Given that the worker will be promised the same utility in both cases and given that the process on x and z are independent we can write the probability of each history h^t as the the product on the probability on the history on z and the probability on x

$$\mathcal{J}(x,z,v|\xi_1) = \sum_t \sum_{(x^t,z^t)} \beta^t \left(f(x_t,z_t) - w^t \right) \pi_{x,t}(x^t|x)\pi_{z,t}(z^t|z)\pi_{\delta,t}(\xi_1),$$

where $\pi_{x,t}$ is the productivity process on x generated by Γ_x , $\pi_{z,t}$ is the process on z generated by $g(z,\iota)$, and $\pi_{\delta,t}(\xi_1)$ is the composition of the leaving probabilities $\tilde{p}(x^t, W^t)$ prescribed by the policy ξ_1. We can then compare the following difference:

$$\begin{aligned}\mathcal{J}(x,z_2,v|\xi_1) - \mathcal{J}(x,z_1,v|\xi_1) = \\ \sum_t \sum_{(x^t,z^t)} \beta^t f(x^t,z^t) \left(\pi_{z,t}(z^t|z_2) - \pi_{z,t}(z^t|z_1) \right) \pi_{x,t}(x^t|x)\pi_{\delta t}(\xi_1),\end{aligned}$$

where we finally use the fact that the transition matrix on z is assumed to be monotonic, in which case we get that all future distributions conditional on z_2 will stochastically dominate distributions conditional on z_1. Given the stochastic dominance of $\pi_{z,t}(z^t|z_2)$ over $\pi_{z,t}(z^t|z_2)$ and the monotonicity of $f(x,z)$ in z we get:

$$\forall t, x^t \quad \sum_{z^t} f(x^t,z^t) \left(\pi_{z,t}(z^t|z_2) - \pi_{z,t}(z^t|z_1) \right) \geq 0$$

which gives the result. See Dardanoni (1995) for more on properties of monotonic Markov chains. □

3.B.4 Characterization of the optimal contract

Lemma 23. *For a given (x, z), a higher wage always means higher lifetime utility.*

Proof. This is a direct implication of the concavity of $\mathcal{J}$ and the envelope condition:

$$\frac{\partial \mathcal{J}(x, z, v)}{\partial v} = \frac{1}{u'(w)},$$

and given also the concavity of $u(\cdot)$, we get that w and v_s are always moving in the same direction. □

Theorem 2 (optimal contract)**.** *For each viable match (x, z), independent of the lottery realization, the wage policy is characterized by a **target wage** $w^*(x, z)$, which is increasing in z such that:*

$$\begin{aligned} w_t \leq w^*(x_t, z_t) \quad &\Rightarrow \quad w_t \leq w_{t+1} \leq w^*(x_t, z_t) \quad && \textit{incentive to search less} \\ w_t \geq w^*(x_t, z_t) \quad &\Rightarrow \quad w^*(x_t, z_t) \leq w_{t+1} \leq w_t \quad && \textit{incentive to search more} \end{aligned}$$

where the target wage is characterized by the zero expected profit condition for the firm:

$$\forall x, z \quad \mathbb{E}_{x'z'|xz} \mathcal{J}(x', z', W_{x'z'}) = 0$$

Proof of Lemma 2 . We start again from the list of first order conditions and we want to find a relationship for wage change.

$$\begin{aligned} \mathcal{J}(x, z, V) = \sup_{\pi_i, W_i, W_{ix'z'}} &\sum \pi_i \big(f(x, z) - w_i + \beta \tilde{p}(x, W_i) \mathbb{E} \mathcal{J}(x', z', W_{ix'z'})\big) \\ s.t \quad (\lambda) \quad &0 = \sum_i \pi_i \left(u(w_i) + \tilde{r}(x, W_i)\right) - V, \\ (\gamma_i) \quad &0 = W_i - \mathbb{E} W_{ix'z'}, \\ &\sum \pi_i = 1. \end{aligned}$$

From the envelope theorem and the f.o.c. for the wage, we get that the wage in the current period is given by

$$i = 1, 2 \quad u'(w_i) = \frac{1}{\lambda} = -\left(\frac{\partial \mathcal{J}}{\partial v}(x, z, v)\right)^{-1}.$$

Now that also means that the wage next period in state (x', z') will be given by

$$\frac{1}{u'(w_{ix'z'})} = -\frac{\partial \mathcal{J}}{\partial v}(x', z', W_{ix'z'}).$$

I then look at the first order condition with respect to W_i

$$\pi_i \beta \tilde{p}_v(x, W_i) \mathbb{E} \mathcal{J}(x', z', W_{ix'y'}) + \beta \lambda \pi_i r'(x, W_i) + \pi_i \gamma_i = 0,$$

where I substitute $r'(x, W) = \tilde{p}(x, W)$, derived in Lemma (3.B.2):

$$\pi_i \beta \tilde{p}_v(x, W_i) \mathbb{E} \mathcal{J}(x', z', W_{ix'y'}) + \beta \lambda \pi_i \tilde{p}(x, W_i) + \pi_i \gamma_i = 0.$$

Using the f.o.c. for $W_{ix'z'}$, which is

$$\beta \tilde{p}(x, W_i) \frac{\partial \mathcal{J}}{\partial v}(x', z', W_{ix'y'}) - \gamma_i = 0,$$

I get the following expression:

$$\pi_i \beta \tilde{p}_v(x, W_i) \mathbb{E} \mathcal{J}(x', z', W_{ix'z'}) + \beta \lambda \pi_i \tilde{p}(x, W_i) + \pi_i \beta \tilde{p}(x, W_i) \frac{\partial \mathcal{J}}{\partial v}(x', z', W_{ix'z'}) = 0.$$

Focusing on $p_1(x, W) > 0$ and $\pi_i > 0$ since otherwise, the worker is leaving the current firm and the next period wage is irrelevant, we first rewrite:

$$\frac{\tilde{p}_v(x, W_i)}{\tilde{p}(x, W_i)} \mathbb{E} \mathcal{J}(x', z', W_{ix'z'}) + \lambda + \frac{\partial J}{\partial v}(s', v_{s'}) = 0.$$

I finally use the envelope condition to extract the wage next period from the last term on the right

$$\frac{\tilde{p}_v(x, W_i)}{\tilde{p}(x, W_i)} \mathbb{E} \mathcal{J}(x', z', W_{ix'z'}) = \frac{1}{u'(w_{x'z'})} - \frac{1}{u'(w)},$$

where since $\tilde{p}_v(x, W_i) > 0$ the inverse marginal utility and consequently wages move according to the sign of expected surplus to the firm. This shows that within each realization of the lottery, the wage will move according to expected profit.

Randomizing over increase and decrease: the next step is to investigate if it is ever optimal for the firm to randomize over a wage increase and a wage decrease at the same time. If the lottery is degenerate then the result holds directly. We are left with non-degenerate lotteries. In that case the first order condition with respect to π must be equal to zero (otherwise we are at a corner solution, which is degenerate). Taking the first order condition with respect to π gives:

$$\beta\tilde{p}(x, W_1)\mathbb{E}\mathcal{J}(x', z', W_{1x'z'}) + \lambda\beta\tilde{r}(x, W_1) = \beta\tilde{p}(x, W_2)\mathbb{E}\mathcal{J}(x', z', W_{2x'z'}) + \lambda\beta\tilde{r}(x, W_2),$$

which we can reorder in

$$\beta\tilde{p}(x, W_1)\mathbb{E}\mathcal{J}(x', z', W_{1x'z'}) - \beta\tilde{p}(x, W_2)\mathbb{E}\mathcal{J}(x', z', W_{2x'yz'}) = \lambda\beta\left(\tilde{r}(x, W_2) - \tilde{r}(x, W_1)\right).$$

Now, suppose that the randomization is over two expected profits of opposite sign for the firm where 1 is positive and 2 is negative. The left hand side is then positive. But in that case we know that $W_2 < V < W_1$ because higher wages give higher utilities in all states of the world, and so they do so also in expectation. This gives us that $\tilde{r}(x, W_2) < \tilde{r}(x, W_1)$. Given that λ is equal to inverse marginal utility it is positive. But then the right hand side is negative, so we have a contradiction. So independent of the randomization, the wage will move according to the sign of the expected profit.

Monotonicity in z: the final step is to show that the efficiency wage is increasing in z. We already know that $\mathcal{J}(x, z, V)$ is increasing in z and decreasing and concave in V. Let's consider $z_1 < z_2$ and associated efficiency wage $w^*(x, z_1)$. We want to show that $w^*(x, z_1) < w^*(x, z_2)$. Call ξ_1 the optimal policy starting at state $\mathcal{J}(x, z_1, V_1)$ where V_1 delivers $w^*(x, z_1)$ and using ξ_1 at (x, z_2), the worker receives V_1 and is paid $w^*(x, z_1)$. The firm makes more profit than at z_1 since $f(x, z)$ is increasing in z and $\mathbb{E}\mathcal{J}$ is larger as well. The optimal policy at (x, z_2, V_1) will pay a higher wage than $w^*(x, z_1)$ to trade some output for a longer expected lifespan, but continue to choose positive $\mathbb{E}\mathcal{J}$. So we found a wage $w_3^* \geq w^*(x, z_1)$ such that $\mathbb{E}\mathcal{J}$ is still positive. This last point implies that $w_3^* \leq w^*(x, z_2)$ and concludes. □

3.B.5 From matching function to tightness

I use the following matching function

$$
\begin{aligned}
p(\theta) &= \theta^{\nu} \\
q(\theta) &= p(\theta)/\theta = \theta^{\nu-1}
\end{aligned}
$$

this gives us that

$$p = q^{\frac{\nu}{\nu-1}},$$

and we have the following equilibrium equality for $q(\cdot)$ from the free entry condition:

we end up with

$$p(x, v) = \left(\frac{1}{k_e} J(x, y, z, v)\right)^{\frac{\nu}{1-\nu}}.$$

Now since I am worried about keeping this function sufficiently concave to insure uniqueness of the worker search decision, I use $\nu < 1/2$.

3.B.6 Recursive Lagrangian formulation

Ignoring the lottery for now, we have the following recursive formulation for $\mathcal{J}$

$$
\begin{aligned}
\mathcal{J}(x,z,V) = \sup_{\pi_i, W_i, W_{ix'y'}} & f(x,z) - w_i + \beta\tilde{p}(x,W_i)\mathbb{E}\mathcal{J}(x',z',W_{ix'y'}) \\
s.t \quad (\lambda) \quad & 0 = u(w_i) + \tilde{r}(x,W_i) - V, \\
(\gamma_i) \quad & 0 = W_i - \mathbb{E}W_{ix'z'}.
\end{aligned}
$$

From which we can construct the Pareto problem

$$
\mathcal{P}(x,z,\rho) = \sup_v \mathcal{J}(x,z,v) + \rho v.
$$

Formally, $\mathcal{P}$ is also the Legendre–Fenchel transform of $\mathcal{P}$, see Villani (2003). We seek a recursive formulation. I first substitute the definition of $\mathcal{J}$ and the constraint on λ in $\mathcal{P}$ to get

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) = \sup_{V,w,W,W_{x'z'}} & f(x,z) - w + \beta\tilde{p}(x,W)\mathbb{E}\mathcal{J}(x',z',W_{x'z'}) + \rho V \\
s.t \quad (\lambda) \quad & 0 = u(w_i) + \tilde{r}(x,W) - V, \\
(\gamma) \quad & 0 = W - \mathbb{E}W_{x'z'}.
\end{aligned}
$$

at which point I can substitute in the V constraint:

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) = \sup_{V,w,W,W_{x'z'}} & f(x,z) - w + \beta\tilde{p}(x,W)\mathbb{E}\mathcal{J}(x',z',W_{x'z'}) + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\
s.t \quad (\gamma) \quad & 0 = W - \mathbb{E}W_{x'z'}.
\end{aligned}
$$

then I append the constraint (γ)with weight $\beta\gamma\tilde{p}(x,W)$

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) &= \inf_{\gamma} \sup_{V,w,W,W_{x'z'}} f(x,z) - w + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\
&\quad -\gamma\beta\tilde{p}(x,W)(W - \mathbb{E}W_{x'z'}) \\
&\quad +\beta\tilde{p}(x,W)\mathbb{E}\mathcal{J}(x',z',W_{x'z'})
\end{aligned}
$$

which finally we recombine as

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) &= \inf_{\gamma} \sup_{V,w,W,W_{x'z'}} f(x,z) - w + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\
&\quad -\beta\gamma\tilde{p}(x,W) \\
&\quad +\beta\tilde{p}(x,W)\mathbb{E}\mathcal{J}(x',z',W_{x'z'}) + \gamma\mathbb{E}W_{x'z'}
\end{aligned}
$$

the final step is to move the sup to the right hand side to get:

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) &= \inf_{\gamma} \sup_{w,W} f(x,z) - w + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\
&\quad -\beta\gamma\tilde{p}(x,W) \\
&\quad +\beta\tilde{p}(x,W)\mathbb{E}\left[\sup_{W_{x'z'}} \mathcal{J}(x',z',W_{x'z'}) + \gamma W_{x'z'}\right]
\end{aligned}
$$

where we recognize the expression for $\mathcal{P}$ and so we are left with solving the following saddle point functional equation (SPFE):

$$
\begin{aligned}
\mathcal{P}(x,z,\rho) = \inf_{\gamma} \sup_{w,W} & f(x,z) - w + \rho\left(u(w_i) + \tilde{r}(x,W)\right) \\
& - \beta\gamma\tilde{p}(x,W) + \beta\tilde{p}(x,W)\mathbb{E}\mathcal{P}(x',z',\gamma). \quad \text{(SPFE)}
\end{aligned}
$$

From the solution of this equation we can reconstruct the lifetime utility of the worker, and the profit function of the firm

$$
\begin{aligned}
\mathcal{V}(x,z,\rho) &= \frac{\partial\mathcal{P}}{\partial\rho}(c,z,\rho) \\
\mathcal{J}(x,z,v) &= \mathcal{P}(x,z,\rho^*(x,z,v)) - \rho^*(x,z,v)\cdot v.
\end{aligned}
$$

3.B.7 Notations

Here is a summary of the notations used in the paper:

β is discount factor

$u : \mathbb{R} \to \mathbb{R}$ is utility function

$c : \mathbb{R} \to \mathbb{R}$ is effort function

e is effort level of the worker

w is wage

x is worker productivity

z is match productivity

$f(x, z)$ is output of worker x in match z

κ is search efficiency on the job

η is vacancy cost

θ is market tightness for market (x, v)

v is value a worker will get in a given submarket

V is value promised to the worker when entering a period

W_i is expected value promised to the worker in realization i of the lottery

$W_{x'z'}$ is value promised to the worker in realization (x', z') of the shock

$v_1(x, z, v)$ is the search policy of the worker

$e(x, z, v)$ is the effort policy of the worker

3.C Additional Data information

3.D Model extensions

3.D.1 severance payments

I present here an extended version of the model with side payments when the worker loses his job. The firm is allowed to choose a value g delivered to the worker when he moves to unemployment.

	Construction etc.	Manufacturing	Retail trade	Services
educ1				
σ_f	0.00498	0.00365	0.00298	0.00471
	(0.000336)	(0.000151)	(0.000215)	(0.000376)
σ_w	0.0245	0.0212	0.0197	0.0279
	(0.00297)	(0.00107)	(0.00176)	(0.00228)
firm perc	16.9	14.7	13.2	14.5
educ2				
σ_f	0.0059	0.00321	0.00431	0.00481
	(0.000434)	(0.000164)	(0.000401)	(0.000509)
σ_w	0.024	0.0185	0.02	0.0254
	(0.00302)	(0.00125)	(0.00269)	(0.00334)
firm perc	19.8	14.8	17.7	15.9
educ3				
σ_f	0.00558	0.00225	0.00521	0.00757
	(0.000768)	(0.000122)	(0.000547)	(0.000299)
σ_w	0.0267	0.0187	0.0224	0.0231
	(0.00342)	(0.00124)	(0.00328)	(0.00173)
firm perc	17.3	10.7	18.9	24.6

Table 3.9: Uncertainty at firm level per industry and education group

I start from the recursive form and

$$f(s) - w - g(1-q) + \beta p_1(e,g)\mathbb{E}J(s', v_{s'}) +$$
$$\rho\left(u(w) + r(e,g)\right) - \mu\beta p_1(e,g)(e - \mathbb{E}v_{s'})$$

where

$$r(e,g) = \sup_{v,q} -c(q) + (1-q)\beta\mathbb{E}U(x',g) + q\beta p(v)v + q(1-p(v))\beta e.$$

and so we get

$$\begin{aligned} r_e(e) &= -q^*\beta(1-p^*) = -\beta p_1(e) \\ r_g(e) &= (1-q^*)\beta\mathbb{E}U_g(x',g) \end{aligned}$$

which we can recombine in

$$f(s) - w + \rho(u(w) + r(e)) - \mu\beta p_1(e)e + \beta p_1(e)\mathbb{E}P(s',\mu)$$

and we get 3 FOC: w μ and e and g

$$e = \mathbb{E}P_\rho(s, \mu)$$

$$-\rho r_e(e) - \mu\beta p_e(e, g)e - \mu p_e(e, g) + \beta p_e(e, g)\mathbb{E}P(s', \mu) = 0$$

$$-(1-q) + \rho r_g - \mu\beta p_g e + \beta p_g \mathbb{E}P = 0$$

I should combine the terms in $p'(e)$ to get $\mathbb{E}P - (\rho + \mu)e$

$$(\mu - \rho)p_1(e) = \beta p_1'(e)\mathbb{E}\Pi_1(s', \mu)$$

and we can recombine the equation in g to find the optimal severance package:

$$(1-q)(\beta \underbrace{\mathbb{E}U_g}_{<0} - 1) = -\frac{\beta p_g}{1-q}\mathbb{E}\Pi_1$$

Note on co-authored work

Note on the joint work in Thibaut LamadonâĂŹs thesis “Dynamic Contracts and Labour Market Frictions”.

Chapter 1, “Repeated Games with One-Dimensional Payoffs and Different Discount factors”, is co-authored between Yves Guéron, Thibaut Lamadon and Caroline Thomas and each author contributed equally to the paper.

Chapter 2, “Identifying sorting with on the job search”, is co-authored between Thibaut Lamadon, Jeremy Lise, Costas Meghir and Jean-Marc Robin and each author contributed equally to the paper.

Chapter 3 is single-authored by Thibaut Lamadon.