Université de Montréal

ESSAYS ON SKILLS AND PRODUCTIVITY

Département de sciences économiques
Faculté des arts et sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences économiques

Fatim Diabagaté, 30 Juin 2021.

Université de Montréal

Faculté des Études Supérieures et Postdoctorales

Cette thèse intitulée :

ESSAYS ON SKILLS AND PRODUCTIVITY

présentée par :

Fatim Diabagaté

a été évaluée par un jury composé des personnes suivantes :

Joshua Lewis
(Président-rapporteur)

Baris Kaymak
(Directeur de recherche)

Joao Galinda da Fonseca
(Membre du jury)

Raquel Fonseca
(Examinateur externe)

Thèse acceptée le 23 Août 2021

à mon Père, Yaya

Remerciements
Acknowledgments

J'ai une profonde gratitude envers mon directeur de recherche Baris Kaymak, pour son soutien indéfectible. Etant parent-étudiante, mon rythme de travail était assez irrégulier ; je lui suis donc particulièrement reconnaissante pour sa patience et sa flexibilité. Merci infiniment de m'avoir accompagnée tout au long de cette expérience. J'ai beaucoup appris de vous.

Je remercie profondément Joshua Lewis, Andriana Bellou et Guillaume Sublet pour leur précieux appui et encouragements. J'aimerais également exprimer ma gratitude envers Rachel Heath, pour m'avoir donné l'opportunité de réaliser une visite de recherche à l'Université de Washington et pour son soutien exceptionnel. Merci également à Fabian Lange pour ses précieux commentaires. Je suis reconnaissante à l'ensemble des professeurs et à l'administration du département de sciences économiques de l'Université de Montréal, ainsi qu'au CIREQ. Je pense particulierement à Marine Carrasco, Benoit Perron, Joao Galindo Da Fonseca et Josée Lafontaine.

J'ai une pensée pour mes collègues qui ont rendu cette expérience à l' Université de Montréal inoubliable. Je pense entre autres à Lucienne, Ngolo, Marlène, Magnim, Guy Leonel, Idriss, et Samuel.

Merci à toi papa de m'avoir motivé à chaque étape de mon parcours académique; à toi maman, à vous mes frères et soeurs pour vos encouragements. Merci à mon époux pour son soutien inconditionnel et, à Seraya, ma fille, grâce à qui j'ai toujours gardé le sourire.

Résumé

Pourquoi devrions-nous vous embaucher? On s'attend presque toujours à cette question lorsqu' on doit passer un entretien d'emploi. Cependant, il n'est pas évident d'établir une réponse standard à cette question. Une formule qui est souvent conseillée est de partager des expériences qui mettent en avant les compétences pertinentes au poste proposé. Pas besoin de préciser que chaque demandeur d'emploi est libre de raconter ce qu'il veut.
Les compétences, voici ce qui intéressent principalement les employeurs. Même si un grade éducatif minimal peut être requis pour certaines positions, ce qui importe, c'est d'être à même d'exécuter les tâches associées au poste proposé. Les systèmes éducatifs sont justement constitués de sorte à développer les compétences que les employeurs recherchent. Déterminer si ces compétences ont effectivement été acquises au sortir de ces cycles éducatifs est une autre question.

Ma recherche utilise des méthodes d'analyse théorique et empirique, macroéconomique et microéconomique, pour examiner diverses problématiques relatives aux compétences. J'aborde la notion des compétences sous différents angles et dans différents contextes: les compétences productives non observables sur le marché du travail, ainsi que les compétences différenciées entre immigrants et natifs. Cette thèse compile deux essais en économie du travail : l'objectif est de contribuer à comprendre comment les compétences peuvent affecter le bien être des agents économiques, mais aussi leurs choix, que ces compétences soient observables ou non. Le premier chapitre aborde la question de la transparence des compétences productives. Lorsque les compétences productives des demandeurs d'emploi ne sont pas observées par les employeurs, un apprentissage de celles-ci s'opère au fil de l'expérience que l'employé acquiert; ce processus d'apprentissage affecte la mobilité des travailleurs et plus particulièrement les risques de mise à pied auxquels ils font face. Dans le second chapitre, les compétences s'assimilent à des avantages comparatifs. Cette section s'intéresse plus spécifiquement à l'impact de l'entrée d'immigrants sur les finances publiques du pays d'accueil. La conclusion est que les compétences relatives des immigrants constituent le facteur ayant l'impact marginal le plus important.

Mots-clés: Compétences, Mobilité professionnelle, Incertitude, Premier emploi, Analyse de survie, Immigration, Politique Fiscale, Dette soutenable.

Abstract

Why should we hire you? This question is almost always expected when going for a job interview. However, it is not obvious to agree on a standard answer to it. One formula that is often recommended is to share experiences that highlight the skills relevant to the position offered. No need to say that each job seeker is free to say what he wants.

Skills are what employers are after. Although a minimum educational degree may be required for some positions, what matters is to be able to perform the duties associated with the proposed position. Education systems are built precisely to develop the skills employers are looking for. Determining whether these skills were actually acquired at the end of these educational studies is another question.

My research uses theoretical and empirical, microeconomics and macroeconomics methods of analysis to examine various economic issues related to skills. I approach the notion of skills from different angles and in different contexts: productive skills not observable on the labor market, and differentiated skills between immigrants and natives. This thesis compiles two essays in labor economics: the objective is to help understand how skills can affect the well-being of economic agents, but also their choices, whether these skills are observable or not. The first chapter addresses the issue of the transparency of productive skills. Productive skills that are not ex ante observed by employers, are learnt throughout workers' experience: this learning process affects the mobility of workers and especially the risk of layoff they face. In the second chapter, skills are equated with comparative advantages. This section looks more specifically at the impact of the entry of immigrants on the public finances of the host country. The conclusion is that the relative labor efficiency of immigrants is the factor with the largest marginal impact.

Keywords: Skills, Labor mobility, Uncertainty, First job, Survival analysis, Immigration, Fiscal policy, Sustainable debt.

Contents

List of Tables

List of Figures

Chapter 1

Employers' Beliefs and Labor Mobility: Insights from First Job

1.1 Introduction

Making the choice to engage into a job relationship is always a risky step. Indeed, such a choice is generally made while relevant information that would make it well-thought is unavailable. As a result, decision making could be cumbersome. And this is true for both parties at play: workers and employers. An employer might offer a job to an applicant, even though some uncertainty about how well the latter will perform in doing it, remains. A worker might accept a job offer, even though some uncertainty about the actual working conditions and workload this job implies, remains.

In this paper, I investigate the implications of this information discrepancy from the perspective of the employer. **I propose a model in three periods to understand the outcome of the first job when workers' actual skills are unobservable**. The employer learning theory (EL thereafter), by Farber and Gibbons (1996), addresses how labor market outcomes should respond to the uncertainty surrounding workers' productive skills if firms behave rationally. Behaving rationally means that firms use all relevant information available. In such a context, what would the first job experience look like? This is the focus of the paper. I specifically study exit out of the first job after schooling is completed, in a model of uncertainty and learning. The model delivers predictions on quits and layoffs. I then provide empirical tests of the implications of the model using US data. Although the model does not cover it, I perform the same estimations on job finding; the results obtained helped substantiate some of the conjectures made over workers' behavior.

As workers start looking for a first job, they are almost strangers to the market, at least when it comes to their productive skills. However some information, as what is found on a CV, can be used to build a prior belief about their abilities.[1] As workers accumulate experience, the market observes measures of their performance, updates the prior belief and ultimately learns workers' actual skills. Most papers have looked at the implications of this process on workers' compensation, showing that workers' pay is explained by observed characteristics like schooling at labor market entry, and progressively by hidden characteristics as experience increases, with the prior predictors playing a marginal role. The argument is that learning has happened throughout experience. While existing tests of this story have extensively focused on wages, less is known about other labor outcomes.

It is fair to consider that the uncertainty surrounding working abilities is at its highest for fresh graduates. So, looking at first job outcomes is the best experiment to study the issues that arise from such information frictions. **Accordingly, in this paper, I focus on the professional experience that started the earliest after the individual left school for good: this is what I refer to as first job**. The analysis provided here is restricted to first job for two additional reasons. Primarily, first job holders constitute the most homogeneous group of workers one can expect to have, when it comes to observed productive characteristics. In fact, as workers evolve on the labor market and move from jobs to jobs, they take divergent trajectories releasing additional signals to the market.[2] Considering multiple jobs would imply to track these information for each worker. Besides, initial professional experiences can shape the overall career making first job especially relevant in such a study on labor mobility. I

[1]Throughout the chapter, I use skill, ability and productivity interchangeably.

[2]As illustration, studies such as Gibbons and Katz (1991) show that a layoff conveys a bad signal to outside employers.

therefore concentrate on first job. The EL theory assumes that relevant information to assess workers' productivity are missing. Although traditional economic theories assume perfect information, modern economics acknowledge that frictions to information exist and cause market failures. Because labor is an essential part of people life, studying how mobility happens in the more realistic context of information incompleteness is worthy. Moreover, a study by Gibbons and Katz (1991) shows that in a context of information incompleteness, if information is moreover asymmetric among employers, a layoff conveys a bad signal to outside employers. A layoff might therefore be detrimental to workers' subsequent opportunities. For layoffs can affect workers' career, this study delivers insightful results to assess workers' labor outcomes. Quits are also interesting to study in a model of employer learning as it answers to whether workers' initiated moves only mirror how opportunities change as the market gets to learn them better, or whether an underlying optimizing behaviour drives their decision over job transition.

Building on the underpinnings of the EL theory, I study a three-period model relying on Kahn (2013) two-period model with asymmetric learning. Those underpinnings are basically that at least a component of workers' true productivity is not observed, and that firms behave rationally. So at baseline, workers are valued based on observed predictors of these unobserved components. As in the seminal paper, the unobserved component I focus on is cognitive skills, and the predictor I consider is schooling. Assuming a certain degree of asymmetry in learning and a random workplace appreciation delivers a consistent setup to distinguish between quits and layoffs; this is a crucial feature because a job spell must terminate either by a quit or a layoff. Learning is asymmetric when incumbent firms and outside firms acquire new information about the workers at different pace. I consider two versions of the model to capture how the asymmetry evolves over time. In the first one, the asymmetry is increasing with full asymmetry as the limit: this information context approximates the short/medium run, where the information advantage of incumbents accumulates. In the second one, the asymmetry is decreasing, with full symmetry as the limit: in this case, I am approximating the long run where the market figures out the ability of everyone.

When it comes to studying variables that are the result of a duration process, survival analysis is generally used. Because one objective of the paper is to perform an empirical test of the model's predictions, I focus on predicting quits and layoffs hazard rates from the model. To identify if EL happens as proposed by Farber and Gibbons (1996), I study how both the hidden skill and its predictor affect layoffs and quits, at the beginning of employment, when labor experience is near zero. I also study the change of those effects over experience. Both versions of the model predict that when experience is short, and ability fixed, increasing schooling raises both the layoff and quit hazard rates. This implies that mobility is high among educated people at the beginning of their career. Indeed, when ability is fixed, increasing schooling over a certain level is similar to being over-educated. Because of asymmetric learning, incumbent firms find out that relying on schooling was overstating the actual productivity, while outside firms do not learn or learn it with a delay. Consequently such workers have a high market value, forcing incumbents that would want to keep them to pay them a prohibitely high wage. Because firms are profit maximizers, these workers face a higher risk to be laid off as a result. Also because they have a high market value, these workers have a higher incentive to quit. In the timing of events, workers can choose to quit only if they receive an effective wage offer from their incumbents, which only happens if their incumbents decide not to lay them off. The positive effect of schooling becomes stronger with time if the information asymmetry between the incumbent and the market widens. However, if the asymmetry is decreasing, which happens in the long run, this positive effect on quits becomes weaker. The change in the effect is not identified for layoffs, because fewer layoffs happen over time, with no layoffs at the limit.

On the contrary, both versions of the model predict that when experience is short, and schooling fixed, increasing ability lowers the layoff and quit hazard rates. This means that mobility is low among workers at the upper tail of the productivity distribution. Indeed, for two workers with same schooling but different ability, the productivity of the one with higher ability will be underestimated. The information advantage of incumbents make them learn this quicker so that the worker with higher ability faces a lower risk to be laid off. Also, because incumbents find out that the worker is good earlier, they are willing to outperform outside offers; thus they offer high compensation, making the worker with high ability less willing to quit. Not surprisingly, the negative effect of ability on turnover widens with time in the case of increasing asymmetry. In the case of decreasing asymmetry, which approximates the long run, the independent effect of ability shrinks, because quits become purely random. The change in the effect is not well identified for layoffs, because fewer layoffs happen over time, with no layoffs at the limit.

The empirical exercises are performed on the random sample of NLSY79, following the literature on employer learning as in Altonji and Pierret (1997). The data provides a measure on cognitive skills that I also use to capture ability. Most results on layoffs that occurred on the first job are supportive of asymmetric employer learning. For workers with up to 4 years of experience, ability affects the layoff hazard rate only marginally, when schooling is fixed. For workers with 4 to 6 years of experience, a one standard deviation increase in ability makes the worker 46% less likely to be laid off and 70% less likely for workers with 6 to 8 years experience, both effects measured while schooling is controlled for. These features of the data are consistent with the hypothesis of increasing asymmetry. As for schooling, during the first 2 years of experience, holding ability constant, one additional year of schooling, makes the workers 22% less likely to be laid off. For workers with 2 to 4 years, this effect decreases to 13%. When experience is larger than 4 years, the effect of schooling is positive and increases with time; however, estimates are relatively imprecise. These results concerning the effect of schooling are consistent with the case of decreasing asymmetry if early signals of productivity received by the current employer are very noisy. Indeed, incumbents will continue to rely heavily on schooling if early performance measures are imprecise; this would explain why schooling reduces the layoff risk for up to 4 years of experience.

Patterns on quit rates yield mixed results. The estimates for newly hired workers are especially consistent with both versions of the model: indeed, the quit rates depend positively on schooling and negatively on ability. Among senior workers, estimations show that schooling is less relevant to predict the quit behavior: this is in accordance with random quits that the model with decreasing asymmetry predicts. Indeed, in the long run, the market should have learned ability of everyone, so that turnover is purely random. Although imprecisely estimated, the results suggest that ability increases the quit hazard rate of senior workers, which neither versions of the model predict.

I also explore the roles of ability and schooling in job finding rates as workers enter the labor market. Some of the results found from the estimation of job finding rates provide alternative explanations that help understand the quit propensity of senior workers at the higher tail of the productivity distribution. Indeed, ability is negatively related to job finding (up to 15% less likely) when unemployment spell is short. Combining these patterns with the quit behavior of senior workers suggests that search effort could also matter. Indeed, the results are consistent with the hypothesis that workers with high ability initially put less effort in finding a job (for instance because they have occupation preferences). Pursuing with the idea of occupation preference,it is plausible that productive workers that gained experience would eventually quit if the first job was not their preferred choice.

There is a large body of work on the topic of employer learning. Indeed, related researches explore various appealing questions. The seminal paper by Farber and Gibbons (1996) introduces the hypothesis of employer learning. Using US data (NLSY), Altonji and Pierret (1997) showed that employers statistically discriminate workers on the basis of schooling and ultimately reward them on the basis of time-invariant skill unobserved by the market. Taking advantage from access to a covariate of ability not observed by employers (AFQT test score)they tested the EL theory on wages' adjustments over the career. They showed that wage dispersion is explained by schooling at early working life, and by hidden and fixed productive skills at late working life. Additional papers used the predictions on wages to either test, refine or challenge the employer learning story. While the basic EL theory assumes the whole market learns workers' ability at the same pace, Schönberg (2007), tests for the hypothesis of asymmetric learning between the current employer and the rest of the market. Learning is said to be asymmetric if the incumbent employer have superior information compared to prospective employers. Schönberg (2007) concludes it is symmetric overall, except for college graduates, potentially. Using a different testing methodology, Kahn (2013) conclude that learning is largely asymmetric. The same exercise was performed by Galindo-Rueda (1993) on British data and conclude that the learning process, especially for blue-collar workers, is asymmetric. On the minus side, a test of EL theory on wages in Germany, by Bauer and Haisken-DeNew

(2001), did not find evidence that employer learning drives wage evolution, except for blue-collar workers at the lower end of the wage distribution. Some arguments provided by the authors to explain this difference with similar studies on US data relates to variability in the quality of schools and universities which is less important in Germany than in the US. The authors argued that "the German apprenticeship system provides standardized occupational training. Therefore, schooling degrees and grades immediately provide more accurate information on the true productivity of an individual than in the US." Indeed, the extent to which hidden factors will increasingly determine wages depends on the accuracy of employers' prior information. Other papers use Farber and Gibbons (1996) test to explore interesting topics. Pinkston (2006) looks at screening discrimination which happens when the market have limited information about one group of workers relative to another. If this is the case, the learning process should be more pronounced for the former. The results on wages of black and white men suggest that the influence of hidden productive characteristics increases faster with experience for black men than for white men, suggesting screening discrimination does happen. Using results on wages, Lange (2007) estimated the speed at which employers learn. The results suggest that employers learn fast. Other papers take a critical stand and question the conclusion that wage adjustments mirror employers' learning process. Kaymak (2014) proposes a different model featuring the same predictions as EL theory. Using NLSY, the results imply that the assumptions required by the alternative model to deliver the results on wages are too strong. Kahn and Lange (2014) proposes a model that nests both learning and human capital hypothesis to explain wage evolution over career. Using an insightful database, they conclude that employer learning comes at play in wage evolution.

The employer learning theory can be seen as a convergence process towards the optimality implied by perfect information in job relationship. The existing papers tested it on wage adjustments over workers' career, assuming either symmetric or asymmetric learning. As of papers that questioned the EL story, they generally conclude that EL hypothesis is important to explain wage dynamics. However, workers' performance on the labor is not restricted to compensation. Job transitions are also important, especially hires and separations initiated by the employer. **Relative to**

the literature, I study the implications of observed and unobserved characteristics of workers on job separations when employers are learning, and when the learning is asymmetric between current and potential employers. Thus, this study of first job contributes to the EL literature by completing the picture on labor market implications of uncertainty and learning. It focuses on outcomes that are really important for workers' career, not only because of the importance of first job but also because of the relevance of the outcomes targeted. The empirical exercises therefore provide a compelling test of the learning hypothesis. Compelling because, if any learning of characteristics that are relevant to assess workers' productivity happens, it should affect hires and layoffs.

The remaining of the paper is structured as follows. In Section 2, I propose a model in three periods and derive the implications for job turnover. Section 3 outlines the empirical strategy to test the predictions of the model before expanding the results obtained. Section 4 concludes.

1.2 A three-period model: first job exit with employer learning

In this section, I study a model in three periods, to derive predictions on first job exit, in a context of uncertainty about workers' productive characteristics and asymmetric employer learning. I build on Kahn (2013) who proposes a two-period model to test for asymmetric employer learning. I add one additional period in the timing, introduce schooling as a predictor of ability and allow for layoffs in addition to quits, to have testable implications on first-job turnover.

Agents and Market. The economy is populated with workers and multiple firms that are ex-ante homogeneous. Information asymmetries create ex-post heterogeneity among firms. Indeed, when focusing on a specific worker, firms are split in two different groups: the incumbent or hiring firm and the outside firms. A worker is characterised by a level of ability (z) and a level of schooling (s), which are fixed over time. Ability and schooling are jointly normal distributed and exogenously given, with the normal conditional distribution of ability $z|s \sim N(\alpha s, \sigma^2)$. Workers' schooling is public information for all agents, but neither the firm, nor the worker observes ability initially. Learning of workers' ability happens but with asymmetries: incumbent firms have an information advantage over outside firms. I consider two different setups to capture these asymmetries in information. The labor market is fundamentally competitive. Firms are risk neutral and workers are risk adverse; both have a time discount factor of 1. As in Kahn(2013), output equals ability in all periods and all firms, so there is no match quality nor human capital accumulation, $y_t = z$.

Timing. There are three periods of employment defined by four times $t \in 0, 1, 2, 3$. At t=0, workers are randomly matched to employers and hired. At the end of the first and second periods of employment, which coincide respectively with t=1 and t=2, firms make wage offers and layoff decisions to maximize expected profits while workers make quit decisions to maximize expected utility from job. At t=3, workers retire.

Information structure. During the first employment period, the incumbent or hiring firm observes a signal of the ability of the worker $\tilde{z} = z + \epsilon$, where $\epsilon \sim N(0, r\sigma^2)$, $0 \leq r$, $z \perp \epsilon$. From this signal, the productivity of the worker is updated to $\tilde{y_1} = y|\tilde{z}, s = z|\tilde{z}, s$: it is the posterior on ability conditional on the realized signal $\tilde{z}$ and s. Using Bayes' rule, $\tilde{y_1} \sim N(\alpha s X + \tilde{z}(1-X), X\sigma^2)$, with $X = \frac{r}{1+r}$.
During the second period of employment, ability is perfectly revealed to the incumbent firm. So $\tilde{y_2} = z|z, s = z$.
I consider two setups to model information asymmetries between incumbent and outside firms. In both of them, the incumbent firm has access to the exact same information. What differs is the information the outside firms have access to during the second employment period.

Information asymmetries. In the first setup (case 1), the information set of outside firms is restricted to schooling s at all periods: they don't observe any additional information. The case 1 is comparable to the short-medium run, where the information advantage of the incumbent accumulates. I therefore refer to the first information context as the case of increasing asymmetry.
In the second setup (case 2), the outside firms do not observe the signal $\tilde{z}$ revealed to the incumbent during the first period of employment; however, the former do observe ability z during the second employment period as the incumbent firm does. The case 2 is comparable to the long run run, where the market figures out ability of everyone. I refer to this second information context as the case of decreasing asymmetry.
Let I_t^j denote the set of information available to the firm j={incumbent,outside} at time t.

The game. The figure below summarizes the timing of events. The first employment period starts at t=0: no action is taken by any of the agents because firms and workers are randomly matched to each other and hiring is granted.

The game starts at the beginning of the second employment period, at t=1. At t=1, the firms engage in a simultaneous game.

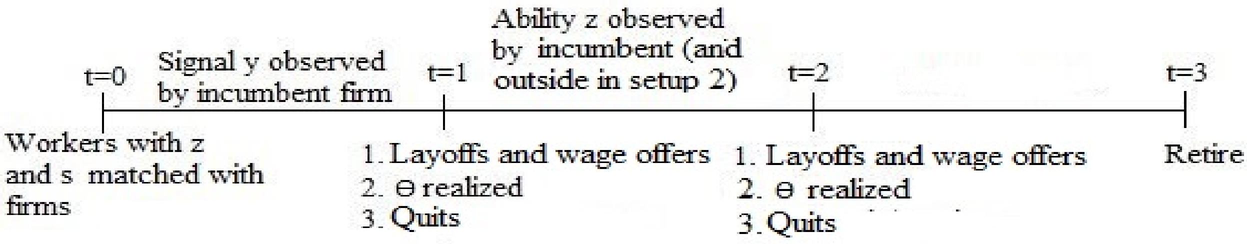


Figure 1.1: Timeline

Both incumbent and outside firms first evaluate the optimal wage they would offer to each worker: they do so based on the information they have access to. Taking the worker as the reference, let's denote $w_t(I_t^{incumbent})$, the optimal wage when the firm is in the incumbent's position, and $v_t(I_t^{outside})$ when taking the role of an outside firm. Following Kahn(2013), I assume that outside firms cannot observe incumbents' optimal wages, otherwise incumbents' private information would be inferred.

At t=1, the incumbent firms also make layoffs' decisions. In fact, once firms have assessed the optimal wage that shall be offered to each worker, they compare the profit expected over all potential workers. At this moment only and not before, optimal wages become effective wage offers only for workers who qualify for employment in the next period: wage offers are extended to workers that feature the highest profit. If the highest profit is reached over the worker currently hired by the firm itself (the inside worker), the wage $w_t(I_t^{incumbent})$ becomes effective; otherwise the inside worker is laid off and the wage offer $v_t(I_t^{incumbent})$ is extended to the outside worker winner of the contest of expected profits.

At this point, workers eligible for employment in the following period have received wage offers. As in Kahn(2013), workers' initiated moves are driven by these wage offers (from incumbent and outside firms) and a random disutility shock θ_{it} which workers learn after each period of employment. This shock comes from the ex post-evaluation of the workplace by the worker i: the θ_{it} are independent and identically distributed with $\theta_{it} \sim U[-\bar{\theta}, \bar{\theta}]$.[3] As the shock θ is realized, workers adjust

[3]Kahn(2013) considers that the θ-shock does not change over time, it is received during period 1. The assumption of a time-dependent θ captures the fact that the evaluation of the workplace is

the incumbent's wage offer and compare net utility of pursuing with the incumbent to expected net utility of accepting the outside offer: the latter coincides with the outside wage offer since the disutility shock is null in expectation. The worker quits if the expected utility from outside is the highest.

Matches that remain active move to the second period of employment. The same game described at t=1 takes place at t=2, which closes the second period of employment. Firms decide which workers are eligible for employment in the last period; they make layoff decisions and extend wage offers to high-value workers. Workers form a new evaluation of the workplace and make quit decisions. Time t=2 also opens the last period of employment for jobs that are still open. Workers retire at the end of this period at t=3.

In the next section, I introduce the problem of the firm formally. All firms solve the same program; the distinction made between incumbent and outside firms is necessary when a specific worker i is considered. Indeed, the incumbent firm of the worker i is an outside firm for the worker i', $i \neq i'$.

In the next section, I introduce the problem of the firm formally. All firms solve the same program; the distinction made between incumbent and outside firms is necessary when a specific worker i is considered. Indeed, the incumbent firm of the worker i is an outside firm for the worker i', $i \neq i'$.

1.2.1 The problem of the firm

At t=0 no decision is made by any of the agents: workers are randomly matched to firms and hiring is granted.

As mentioned above, agents start acting at t=1, and decisions are made at the beginning of the second and last employment periods, coinciding with t=1, 2, respectively. At t=1, 2, the firm solves two problems sequentially. Firstly, the employer evaluates what wage to offer to all workers, both inside ($w_t(I_t^{incumbent})$) and outside

random and subject to changes based on unpredictable conditions.

($v_t(I_t^{outside})$); these wages are set to maximize expected profit over each of them. The expected profit is equal to the expected output minus the wage to be offered. Secondly, expected profits are compared and wage offers are extended only to workers featuring the maximum profit. For a specific firm j hiring a worker i, if the maximum output is reached over an outside worker i', then the worker i is laid off.

The labor market is competitive at t=0, with information perfectly symmetric among firms. Wages equal expected productivity conditional on all relevant information. Since the only information available to firms at t=0 is workers' schooling, $w_0 = E(z|s)$.

At t=1 and 2, information asymmetries create rents. At t=1, which coincides with the end of the first employment period and the beginning of second one, the firm evaluates the candidate wage that shall be offered to each of the workers for the coming employment period. For inside workers, the objective of the firm is to maximize its profit over the two remaining periods of employment. Thus, the program of the firm over inside workers is solved recursively, starting with t=2 which opens the last period of employment. For outside workers, I assume firms offer spot contracts.[4] For now on, I will refer to incumbent worker, when I discuss the program the worker currently hired by the firm, and to outside worker when I consider workers hired elsewhere.

So optimal wages maximize firms' expected profit: in the case of the incumbent worker, the expected profit is computed over all remaining periods while it is restricted to very next period for outside workers. In addition to the fact that actual productivity is not yet revealed at t=1, the randomness of the profit also arises from the worker ex-post evaluation of the workplace, not revealed when the wage offers are made. Therefore, the workers may or may not accept the wage offer. The analytical expression of the expected profit incorporates the probability that the worker accepts the wage proposal; I discuss this probability first.

Probability that the worker accepts the wage offer of the incumbent. The worker ac-

[4]The reason is that I focus on workers' first job. What happens to workers that quit as they move to a new job is out of the scope of this study.

cepts the wage offer of the incumbent if the net utility from inside ($w_t - \theta_t$) exceeds the outside wage offer (v_t). So the probability that the worker accepts the incumbent's wage offer writes as:

$$P_{accepts}(t) = P(w_t - \theta_t \geq v_t) = P(\theta_t \leq w_t - v_t) \tag{1.1}$$

To simplify the writing, w_t is the wage as an incumbent worker $- w_t = w_t(I_t^{incumbent})$, and v_t the wage as an outside worker $- v_t = v_t(I_t^{outside})$; θ_t is uniformly distributed on $[-\bar{\theta}; \bar{\theta}]$ and captures the worker's workplace appreciation. The lower the disutility shock θ_t is, the better the workplace appreciation, and more chance there is that the worker accepts the incumbent wage offer.

Programs at t=2. At this point the firm knows the true productivity y=z of the incumbent worker, but the net utility that worker draws from the job is not yet revealed. As firms derive optimal wage for incumbent workers, their objective is to maximize the expected profit over the possible realizations of θ. Formally, the problem of the firm relative to the incumbent worker, is to choose w_2 that maximizes the value of the current match V_2:

$$V_2(z) = max_{w_2}(z - w_2)P_{accepts}(2) \tag{1.2}$$

$$V_2(z) = max_{w_2}(z - w_2)P(\theta_2 \leq w_2 - v_2) \tag{1.3}$$

While w_2 if the choice variable of the firm here, v_2 is the outside option of the incumbent worker, namely, the wage the worker would be offered by prospective employers outside. Productivity of incumbent workers is perfectly observed, so output equals ability. Because incumbent firms know the information available to outside firms they can infer what would be their incumbent workers' outside options, v_2. With this information in hand, they can assess the probability that the worker accepts the wage offer they would eventually make. It appears that the value of the incumbent worker at t=2 is a positive function of the worker's productivity, which is pretty straight. However, it is a concave function of w_2 due to two forces at work. First, the profit itself is a negative function of w_2. Second, the probability that the workers accepts the wage offer is a positive function of w_2. So, an optimal wage exists over which the value of the worker takes a decreasing trend:

$$w_2^*(z,s) = \frac{1}{2}(z + v_2 - \bar{\theta}) \tag{1.4}$$

Leaving the last term aside for now, we note that the optimal wage at t=2 is an average of true productivity z and the worker's outside option v_2.[5] The rationale behind this is economically sound. Since firms have learned true productivity of their incumbent workers, they hold an information advantage over them and over outside firms also. These information advantages may or may not translate into rents.

The value of the information rent over their workers is $z - w_2^*(z,s)$. Indeed, firms could set their potential offers to the their worker's outside option plus a marginal amount, $v_2 + \epsilon$, $\epsilon > 0$; they would thus grab almost all the information rent. What impedes them to do so is the random disutility shock that is not revealed at the time firms are solving for optimal wages. In fact, by setting their pay proposal to $v_2 + \epsilon$, they face the risk of losing workers who have a high realization of θ_2, the disutility shock. Alternatively, if the firm sets a wage proposal equal to z, there is a bigger chance of outweighing θ_2 for high quality workers. By doing so, firms are more likely to keep high quality workers, at the cost of loosing the entire information rent. The optimal choice appears to be the alternative in between. Indeed, to keep high quality workers and capture some of the value of the information rent at the same time, firms set pay to their incumbent workers equals to the average of the two competing choices, z and v_2.

The same discussion can be done concerning the information advantage over outside firms. The main difference it that the rent over outside firms is fully realized as soon as firms have access to their private information: in fact, the rent is completely determined by the sign of $z - v_2^*$. If $z > v_2^*$, then $z > \frac{1}{2}(z + v_2^*)$. Consequently, firms benefit from observing z privately. On the contrary, if $z < v_2^*$, implying that $z < \frac{1}{2}(z + v_2^*)$, firms are penalized by the private information they have on their incumbent workers.[6]

[5] See Appendix .1.1 for proof.

[6] It is not obvious that all workers featuring $z < v_2^*$ would be laid off, although most of them would be; this would be clearer as I discuss firms' second problem.

Optimal wages of incumbent workers are linked to true ability, z, even though neither them, nor outside firms can observe it. They are furthermore increasing in z; so the larger z is, the more likely the incumbent worker would accept the insider wage once extended. Thus, as in the two-period model in Kahn(2013), the three-period version also features the classical lemons effect (a la Greenwald 1986) that workers at the lower tail of the skill distribution are more likely to quit.

The last term $-\bar{\theta}/2$ originates from the randomness of turnovers. Indeed, the uncertainty linked to workers' satisfaction from the job creates a risk. Because firms are risk neutral and workers risk adverse, this is a direct source of rent over incumbent workers. Specifically, this last term captures a risk premium charged to the incumbent worker. The reason is that the firm will provide insurance against the risk associated with the disutility shock if the incumbent worker happens to feature the highest profit. In fact, incumbent workers that qualify for employment in the next period would be offered a risk-free wage before the disutility shock is revealed. This wage is equal to the optimal wage w_1^*, which will be binding upon realization of the shock that captures job satisfaction. Yet, while extending the wage offer, the firm is still uncertain as to whether the incumbent worker decides to quit, which will happen if the disutility shock is too high.

Let's turn now to outside workers. Because they are offered spot contract, the optimal pay, v_2^*, equals expected ability conditional on all the information available to the firm: $I_2^{outside}$ and the willingness of the worker to reject the wage that the incumbent would offer.[7]

$$v_2^* = E(z|I_2^{outside}, rejects\ \mathrm{w}_2^*)(1.5)$$

For outside workers, as firms condition on the event that workers reject the wage offer from their incumbent employers, w_2 is evaluated in expectation. Indeed they do not observe w_2 otherwise their information disadvantage would be virtually inexistent.

[7]Depending on the degree of information asymmetry, as captured by the two setups I consider, $I_2^{outside}$ is either {s}, or {s,z}

Now that potential wage offers to both incumbent and outside workers are set, the firm compares the all profits expected over all prospective employees. The profit expected over the incumbent worker is ($V_2(z)$), while the profit expected over the outside workers is null ($E(z|I_2^{outside}, rejects) - v_2$ = 0). The wage offer becomes effective for workers that feature the highest value. Typically,

$$w_2^* \text{ is effective if } V_2(z) \geq 0 \tag{1.6}$$

$$v_2^* \text{ is effective if } V_2(z) < 0 \tag{1.7}$$

If the second condition holds, the incumbent worker is laid off. Otherwise the job remains active and the wage offer w_2 is extended to the incumbent worker.

Programs at time t=1. Moving up to t=1, agents make their first move: firms evaluate wages that would maximize the profit over each worker. Time t=1 closes the first period of employment and the information set on incumbent workers comprises schooling (s) and a signal of ability $\tilde{z}$ received during the first employment period. The objective of firms here, is to find the wage w_1 that maximizes the expected profit over the two remaining periods of production.

$$V_1(\tilde{z}) = max_{w_1}[z - w_1 + E(V_2(z))|\tilde{z}, s]P_{accepts}(1) \tag{1.8}$$

$$V_1(\tilde{z}) = max_{w_1}[\tilde{y_1} - w_1 + E(V_2(z))|\tilde{z}, s]P(\theta_1 \leq (w_1 - v_1))$$

As seen before, the output expected over the incumbent worker is $\tilde{y_1} = E(z|s, \tilde{z})$, the posterior on ability conditional on schooling s and the realized signal $\tilde{z}$. It appears that the value at t=1 is a positive function of the profit expected for t=2. Let's precise that V_1 is the value of the worker over the two remaining periods of employment, evaluated at t=1. Although the firm does not discount the future, in assessing V_1, a larger weight is given to the profit at t=1 compared to the one given to the profit expected at t=2. The reason is that the profit at t=1 is more likely to get to realization than the profit at t=2, simply because reaching t=2 is conditional on passing time t=1.

The optimal wage writes as:

$$w_1^* = \frac{1}{2}(\tilde{y}_1 + E(V_2(z)|\tilde{z}, s) + v_1 - \bar{\theta}) \tag{1.9}$$

Since firms are rationale forward-looking profit maximizers, they are interested in keeping high value workers, so optimal wage at t=1 is also a positive function of the profit expected for t=2. This is pretty intuitive. Indeed, the larger the profit expected tomorrow, the larger the compensation that will be afforded to incumbent workers today, so that they are incentivized to stay.[8]

The structure of the optimal wage at t=1 is exactly the same as the one discussed at t=2. This is not surprising since firms face the same trade-off: maintaining high value workers while keeping the maximum of the informational rent. Specifically, at t=1, before the wage is set, the total return from keeping the incumbent worker is the sum of the productivity expected over the coming period of employment, $\tilde{y}_1$ (the subscript is 1 because this expectation is formed at t=1) and the value expected over last period of employment, $E(V_2(z)|\tilde{z}, s)$ (the subscript 2 follows the same logic). By setting the wage at t=1 to $\tilde{y}_1 + E(V_2(z)|\tilde{z}, s)$, firms will be loosing the entire information rent to secure workers at the higher tail of the ability distribution. Alternatively, by setting it to the outside option of their incumbent workers (plus a marginal positive amount $v_1 + \epsilon$), the firm would keep most of the information rent at the expense of facing a higher risk to loose productive workers. We saw earlier that the optimal position for the firm is to be right in the middle and give each of these competing alternatives the same probability.

The last term, $-\bar{\theta}/2$ is the risk premium that the firm charges to the incumbent worker to provide insurance against the risk carried by the disutility shock θ_1.

As for outside workers, exactly as explained before, they are offered spot contracts. Thus, the optimal offer v_1^* equals expected ability conditional on all the information available to the firm: schooling (s) and the fact that the worker rejects the wage offer received from inside. At t=1, the information set on outside workers is the

[8] See Appendix .1.2 for proof.

same in either of the two setups.

$$v_1^* = E(z|s, rejects\, w_1^*) \tag{1.10}$$

Similarly to t=2, once these optimal wages are set, the firm compares the profits that each type of worker provides in expectation. However, firms are myopic when they compare expected values at t=1; this is one limitation of the model. Specifically, the comparison is made only over the profit expected for the very next period of employment for both incumbent and outside workers. This is so because I only study first job and subsequent job relationships are out of the scope of the current study. Thus, if an outside worker is finally hired at t=1, the wage that shall be offered at t=2 is not modelled. It would be irrational to compare profits expected over different time horizon. Consequently, at t=1, the optimal wage becomes effective for workers that feature the highest profit at t=1: these workers are eligible for employment in the coming period. Typically,

$$w_1^* \text{ is effective if } E(z|\tilde{z}, s) - w_1^* \geq 0 \tag{1.11}$$

$$v_1^* \text{ is effective if } E(z|\tilde{z}, s) - w_1^* < 0 \tag{1.12}$$

If the second condition is satisfied, the incumbent worker is laid off. Otherwise the job remains active and the wage offer w_1^* is extended to the incumbent worker.

The next section introduces the problem of the worker. For now on, the worker is the reference, I will thus use incumbent firm for the hiring firm and outside firms for all prospective employers present in the market.

1.2.2 The problem of the worker

At t= 0, the worker has no decision to make because matching is random and hiring granted.

At t=1, which opens the second period of employment, the workers eligible for employment during this second period receive wage offers either from the incumbent firm w_1, either from outside firms v_1 or from both incumbent and outside firms. Then, the realized evaluation of the workplace θ_1 is revealed and the worker adjusts the wage offer of the incumbent firm with this disutility shock θ_1.[9] Then, the worker decides whether to accept the incumbent firm's wage offer and pursue the match, or to accept the wage offer from outside and quit. Because I focus on workers' first job, the problem of the worker is static. Indeed, what happens to a quitter in the subsequent periods is out of the scope of this study since this is part of a new job relationship. Typically, the decision rule of the worker at t=1 is the following:

$$Quit \text{ if } w_1 - \theta_1 < v_1 \tag{1.13}$$

$$Stay \text{ if } w_1 - \theta_1 \geq v_1 \tag{1.14}$$

The same process described at t=1 happens at t=2 and the decision rule of the worker at t=2 is:

$$Quit \text{ if } w_2 - \theta_2 < v_2 \tag{1.15}$$

$$Stay \text{ if } w_2 - \theta_2 \geq v_2 \tag{1.16}$$

Now that the strategies of all players have been discussed, I define the equilibrium.

1.2.3 The Equilibrium

The Perfect Bayesian equilibrium (PBE) consists of an incumbent wage offers $w_t^*(I_t^{incumbent})$ at times t=1 and 2, an outside firm wage offers $v_t^*(I_t^{outside})$ at times 1 and 2, a layoff rule for the incumbent firm in periods 1 and 2 and a quit rule for workers in periods 1 and 2, such that: (1) the hiring firm maximizes profits conditional on the behavior of outside firms, (2) outside firms maximize profits subject to beliefs about worker's quit behavior, (3) workers maximize job net return in making their quit decisions

[9] The risk aversion of workers imposes that the utility be concave. However this assumption was only made to explain the rationale behind the risk premium charged by incumbents. Apart from that, the increasing trend of the utility is enough for all the other results. So I go on with the linear utility.

and (4) all beliefs are consistent with these strategies.

In the next section, I will be deriving the implications of the model for job turnover. The objective is to formulate testable predictions that will be confronted to the data. Clearing these predictions will raise the understanding of the mechanisms governing workers' mobility.

1.2.4 Implications of the Model For Job Turnover

The employer learning test proposed by Altonji and Pierret (2001) lies on assessing how a labor market outcome, generally wages in existing studies, is affected by both a hidden component of workers' productivity (ability z here) and a prior predictor of the hidden skill (schooling s here): the way these effects change over time should follow a specific pattern. More specifically, no matter the direction of the effect, the hidden characteristics should matter more over time, so the magnitude of their effect should increase. The prior predictors of these hidden components of productivity should matter less over time, so the magnitude of their effects should decrease over time. The two outcomes I investigate relate to job turnover, and especially layoffs and quits occurring in the first job.

I thus expand the implications of the model on layoffs and quits hazard rates: $\lambda_L(t)$ refers to the layoff hazard rate and $\lambda_Q(t)$ to the quit hazard rate. Because the outcomes of the job depend on the information structure, I will discuss these implications for each of the two configurations described earlier . In the first setup, the asymmetry is increasing with full asymmetry as the limit, and in the second one the asymmetry is decreasing with full symmetry as the limit. I thus discuss the strategies of both the firms and the workers more specifically. Indeed, the precise optimal wages depend on the information that outside firms have access to.

Case 1: Increasing Asymmetry — Information is asymmetric at t=1 and fully asymmetric at t=2 (outside firms never learn ability)

Quits and layoffs rules are fully determined by wage offers and the disutility shock that the worker receives. Since the later is purely random, I introduce the optimal wages first. I would then have the key elements to elaborate on quit and layoff rules.

Optimal wages. The program of incumbent firms is dynamic so their decisions at t=1 depends on what they expected for t=2. I therefore start with t=2. In the setup under study, outside firms only observe schooling throughout all periods of employment. They also condition on the willingness of workers to quit by rejecting their incumbent wage offer. Solving the model with the assumption that outside firms do not acquire additional information yields the following optimal wages:

$$v_2^*(s, rejects) = E(z|s, rejects) = E(z|s) - b_2 \tag{1.17}$$

$$w_2^*(z, s) = \frac{1}{2}(z + v_2^* - \bar{\theta}) = \frac{z + E(z|s) - b_2 - \bar{\theta}}{2} \tag{1.18}$$

As explained before, outside firms offer spot contracts. Thus, optimal outside wage v_2^* equals expected productivity, conditional on all information available. Here, outside firms have not observed any additional information: schooling remains the only information they have access to, at t=2. Still, as they set their optimal offers, outside firms condition on the fact that the worker rejects the incumbent firm's wage offer. When a worker is willing to quit, outside firms infer that the incumbent has made an unattractive wage offer, which means that the incumbent firm must have observed a relatively poor performance of the worker: this comes as a penalty captured by b_2 in the optimal offer of the outside firm.[10]

As seen in the discussion over the program of firms, the optimal compensation scheme of incumbent firms gives equal weights to the two competing choices they are facing: capturing the maximum of the information rent and retaining productive workers for the last period of production. Since these wage offers will be binding,

[10] See Appendix .1.3 for proof.

incumbent firms also charge a risk premium to workers for insuring them against them risk carried by the disutility shock.

Before turning to optimal wages at t=1, w_1^* and v_1^*, since w_1^* depends on $E(V_2|\tilde{z}, s)$ and v_1^* conditions on what w_1^* could be, I first discuss $E(V_2|\tilde{z}, s)$, the value that incumbent firms expect at t=2:

$$E(V_2|\tilde{z}, s) = \frac{1}{4\bar{\theta}}[\overbrace{(E(z|\tilde{z}, s) - E(z|s)}^{\textbf{(1)}} + \underbrace{b_2 + \bar{\theta}}_{\textbf{(2)}})^2 + \overbrace{Var(z|\tilde{z}, s)}^{\textbf{(3)}}] \tag{1.19}$$

Expected value of the worker at t=2 is the sum of three terms:

(1) captures the **information rent** of incumbent firms over outside firms. During the first period of employment, incumbent firms have observed a signal of ability $\tilde{z}$ which is not accessible to outside firms. With this private information, the former corrects initial beliefs over the ability of the worker $E(z|s)$ while the latter cannot do the same. Since (1) + (2) is squared, the larger $E(z|\tilde{z}, s) - E(z|s)$ is, the more room incumbent firms have to make profit while keeping good workers. Since incumbent firms are forming expectations at t=1, this is the best prediction they can make with the information available at t=1.

(2) adds the two **penalties** that firms charge to workers at t=2: b_2 which outside firms deduct on quitters, and $\bar{\theta}$ which captures the fee levied by incumbent firms for insuring workers. The larger b_2 is, the less attractive the outside wage would be, and the more chances there are that the value expected at t=2 comes to realization. Recalling that firms are risk neutral, while workers are risk adverse, the larger $\bar{\theta}$ is, the larger the risk associated with the disutility shock, and the larger the risk premium that incumbent firms charge to their workers for making binding wage offers before the shock is realized.

(3) measures the **precision of the signals** of ability $\tilde{z}, s$. The larger $Var(z|\tilde{z}, s)$, the less reliable the signals $\tilde{z}, s$ are. In fact, workers selected for employment during the second period of production are those who took the lead in the contest of expected profits; thus, since expectations are on average correct, they are on average high productive workers. Therefore, the larger the remaining variability in z, the higher the maximum z can be, so that incumbent firms can hit a very high level of productivity in the last period of production.[11]

At t=1, the information structure is the same in the two information settings. Still, because the program of incumbent firms is dynamic, w_1^* depends on the expected value of the worker at t=2, $E(V_2|\tilde{z}, s)$ and especially on v_2^*. Also, because outside firms condition on workers' rejection of incumbent's offer, v_1^* also depends on what outside firms expect w_1^* is. This yields the following optimal wages at t=1, starting with v_1^*:

$$v_1^*(s, refuses) = E(z|s) - b_1 \tag{1.20}$$

We saw from the discussion at t=2, that outside firms penalize potential quitters because their willingness to move signals that the offer from their incumbent was not appealing enough. Yet, outside firms do know that incumbents have an information advantage over them. As rationale agents, they somehow correct their beliefs learning from actions taken by workers. Let's note that the penalty applied here is different from the one charged at t=2; it is actually larger ($b_1 > b_2$).[12] I first expected the reverse, but thinking through the model, it appears that the more uncertainty there is, the more rent agents can make. At t=2, which closes the second period of employment, incumbent firms have observed workers' actual productivity. Thus there is less uncertainty in the signal conveyed by quit behavior. However, at t=1, incumbent firms observed a noisy signal of true productivity, so quit behavior conveys a mixed signal of true productivity, noise and disutility shock. This extra deduction applied at t=1 is somehow comparable to incumbent firms charging a risk premium to workers for the risk carried by the disutility shock.

As for the optimal wage from incumbent firms, after v_1^* and $E(V_2(z)|\tilde{z}, s)$ are replaced by their expressions, w_1^* writes as:

$$w_1^* = \frac{\frac{(5+\frac{b_2}{\theta})}{4}E(z|\tilde{z}, s) + \frac{(3-\frac{b_2}{\theta})}{4}E(z|s) - (b_1 + \bar{\theta}) - \frac{1}{8\theta}.K}{2} \tag{1.21}$$

[11] See Appendix .1.3 for proof.

[12] See Appendix .1.3 for proof.

$$\text{with } K = (E(z|\tilde{z}, s) - E(z|s))^2 + (b_2 + \bar{\theta})^2]$$

Comparing w_1^* to w_2^*, it is interesting to note that at t=1, the share of the expected output $E(z|\tilde{z}, s)$ conceded by incumbent firms ($\frac{1}{2}(\frac{5}{4} + \frac{b_2}{4\theta})$)is larger than the one granted at t=2 ($\frac{1}{2}$). This result reveals a strategic behaviour of firms which consists in paying high wages in the present to induce productive workers to remain long enough so that firms can catch up in subsequent periods.[13]

Now that optimal wages are known, let's pursue with the decisions of agents related to turnover. In the empirical part, the labor market outcomes I consider are layoff and quit hazard rates.

The layoff (respectively quit) hazard rate at time t is the probability that a layoff (respectively quit) happens at time t, conditional on not exiting (no layoff, nor quit) at t-1. Since all workers are present at t=0, by definition, the layoff (respectively quit) hazard rate at t=1 equals the layoff (respectively quit) probability at t=1. As for the layoff (respectively quit) hazard rate at t=2, it is equal to the layoff (respectively quit) probability at t=2, conditional on surviving time t=1 (no layoff, nor quit at t=1).

Following the timing of events, I start with the second problem of the firm: which worker qualifies for employment in the next period. I then turn to the quit rule of the worker: which wage offer to accept. Once these two events have been discussed, I explicit the survival probability at t=1. I would then have all the ingredients to derive the layoff and quit hazard rates.

The layoff rule of the firm. The first problem the firm consists in finding what wage maximizes the expected profit over each worker, both inside and outside. Then, the firm compares all expected returns from matching with each worker. In the event that the expected profit associated to the worker currently hired is below the outside option of the firm, the worker is laid off. The outside option of the firm is zero because as we saw, outside workers are offered spot contracts and the wage offer made to them is exactly equal to their expected output.

[13] See Appendix .1.3 for proof on w_1.

The worker is laid off at time t if:

$$E(y_t|I_t^{incumbent}) - w_t^* < 0 \tag{1.22}$$

Workers are laid off at t if the output incumbents expect at t is below the optimal wage they would offer. As mentioned earlier, applying this rule at t=1 is a consequence of the myopia of firms imposed by the the focus on the first job spell.

Workers also can initiate a separation by quitting. Following the timing of event, a quit can happen only when the wage offer from the incumbent is effective, which means that laying off the worker was not profitable to the firm.

The quit rule of the worker. The worker compares the return between pursuing with the same employer and taking the outside offer.
The worker i decides to quit at t if:

$$w_t^*(I_t^{incumbent}) - \theta_t < v_t^*(I_t^{outside}) \tag{1.23}$$

Workers move to where their net utility is the highest. The disutility shock is null in expectation so the outside offer coincides with the expected net utility the worker gets by moving (v_t^*)

I can now describe the survival event at t=1. I need this because layoff and quit hazard rates at t=2, are conditional on surviving at t=1.

The condition for job continuity (survival) at t=1. The worker survives in the job past t=1 if neither a layoff nor a quit are profitable at t=1. The job survives if:

$$E(z|\tilde{z}, s) - w_1^* >= 0 \ \& \ w_1^* - \theta_1 >= v_1^* \tag{1.24}$$

Workers that pursue their job past t=1 are those that were not laid off, and chose to remain with their incumbent firm.

All is now set to define layoff and quit hazard rates.

Layoff hazard rates. At t=1, the layoff hazard rate is equal to the layoff probability.

$$\lambda_L(1) = P[E(z|\tilde{z}, s) - w_1^* < 0]$$

$$\lambda_L(1) = P[\frac{1}{2}((\frac{1}{2} - \frac{b_2}{2\bar{\theta}})E(z|\tilde{z}, s) - (\frac{3}{2} + \frac{b_2}{2\bar{\theta}})E(z|s) + b_1 + \bar{\theta} - \frac{1}{4\bar{\theta}}K) < 0] \quad (1.25)$$

$$\text{with } K = B^2 + (b + \bar{\theta})^2 + V(z|\tilde{z}, s)$$

At t=2, the layoff hazard rate is equal to the layoff probability at t=2, conditional on surviving at t=1:

$$\lambda_L(2) = P[\frac{z - E(z|s) + b_2 + \bar{\theta}}{2} < 0|E(z|\tilde{z}, s) - w_1^* >= 0, w_1^* - \theta_1 >= v_1^*] \quad (1.26)$$

Quit hazard rates. At t=1, the quit hazard rate is equal to the probability of a quit conditional on no layoff at t=1:

$$\lambda_Q(1) = P[w_1^* - \theta_1 < v_1^*|E(z|\tilde{z}, s) - w_1^* >= 0] \quad (1.27)$$

At t=2, the quit hazard rate is equal to the quit probability at t=2, conditional on surviving t=1 and no layoff at t=2:

$$\lambda_Q(2) = P[w_2^* - \theta_2 < v_2^*|E(z|\tilde{z}, s) - w_1^* >= 0, w_1^* - \theta_1 >= v_1^*, z - w_2^* >= 0] \quad (1.28)$$

Below, I formulate propositions describing how schooling s and ability z affect the layoff and the quit hazard rates through time. These propositions are hypotheses to be tested in the empirical section.

Proposition 1 *: Employer learning and layoffs — Case of increasing asymmetry*
Holding schooling constant, the layoff hazard rate is negatively related to ability at both t= 1 and t=2, with a larger magnitude at time t=2.

Holding ability constant, the layoff hazard rate is positively related to schooling at both t=1 and t=2, with a larger magnitude at t=2.

Proof 1 *See Appendix .1.3*

In the case of increasing ability, where the information advantage of the incumbent accumulates, the predictions concerning the effect of ability on layoffs are in accordance with what previous studies find for wages, Specifically, high ability is associated with good labor market outcome, and this relationship becomes stronger over time. Not consistent with findings on wages, higher schooling increases the risk to experience a bad labor outcome, and this relationship becomes stronger over time.

To illustrate more, let's compare two workers with same schooling s but different levels of ability z. The worker with higher ability faces a lower risk to be laid off at both t=1 and t=2: this is a good labor outcome, comparable to a wage increase. Besides, the worker with higher ability faces a risk that is even lower at t=2 compared to the risk faced at t=1.

As for the predictions concerning the effect of schooling, it is intriguing to see that schooling increases the layoff risk, when ability is fixed. In fact, because schooling is the public information used to predict ability initially, wage offers at t=0 only depend on schooling. Let's analyse the situation of two workers with the same ability z but different schooling s. The worker with higher schooling will be offered a higher wage at time t=0. During the first period of employment, meaning between t=0 and t=1, incumbent firms acquire additional information on these two workers. Because the hiring firm observes new signals of ability, the expected output will be a convex function of this new signal and schooling. Since these two workers have same ability, the expected outputs will be closer from each other at t=1 than they were at t=0. However, the worker with high s will have an outside wage offer that is higher, because outside firms only observe schooling and workers' decision to move. Therefore, the worker with high s will be offered a higher wage from inside at time t=1 for the incumbent firms to have chances of retaining good workers. However, as discussed earlier, the outputs expected from these two workers at t=1 are close because they have identical ability. This leaves the incumbent firm hiring the high s worker with an expected profit that is lower in expectation. As a result, this worker will face a higher risk to be laid off. At t=2, the worker with high schooling faces a risk that is even higher. Indeed, while expected output at t=1 were just close to each other, they are exactly the same at t=2. Yet, the optimal wage from the incumbent is an average of incumbents' expected productivity — which is z at t=2, thus fully independent of s when z is held constant — and the outside option which depends on schooling only. Thus, the worker with higher s would be offered a higher wage while incumbent is expected the same output from both. Therefore, at t=2, the layoff risk is even higher for workers with high s.

The discussion made above explains why schooling affect layoffs differently from wages. Indeed, it is the case that the optimal wage is positively related to schooling, and less over time. It is also the case that the output that the incumbent expects is positively related to schooling and less over time. However, the wage depends more on schooling than the expected output, simply because incumbents have to consider workers' outside option in their wage setting. Moreover, the loss in predictive power of schooling is larger for expected output than it is for wages because the information asymmetries widens over time: incumbent learn while outside firms do not. Consequently, the profit which is the difference between the expected output and the wage, is negatively related to schooling, and even more over time. Since firms are profit maximizers, their decision concerning turnover depend on what profit they expect and not on how they will compensate workers. This explains why both the sign and the change in the effect of schooling on layoffs are opposite to those predicted for wages.

Proposition 2 *: Employer learning and quits — Case of increasing asymmetry*
Holding schooling constant, the quit hazard rate is negatively related to ability at both t=1 and t=2, with a larger impact at t=2. The magnitude of both effects depends on the distribution of θ. The more variability there is in θ, the lower both effects are.

Holding ability constant, the quit hazard rate is positively related to schooling at both t=1 and t=2 with a larger impact at t=2. The magnitude of both effects depends on the distribution of θ. The more variability there is in θ, the lower both effect are.

Proof 2 *See Appendix .1.3*

As said earlier, the quit decision of workers is static because the model only covers the first employment spell. At each t, workers that qualify for employment in the subsequent periods receive wage offers from both their incumbent and outside firms. Then they find out about the level of satisfaction with the workplace. Once they adjust the incumbent's wage offer with the disutility shock, they compare their outside options with the net return of their current job.

Because outside firms only observe schooling at both t=1 and t=2, the outside options of workers is fully determined by their schooling s at t=1 and t=2. As for the incumbents, since they accumulate information allowing them to fully learn ability, their wage offers depend on both ability and schooling. Because incumbent firms have additional information that is relevant, they rely less on schooling than outside offer do. These points altogether explain the sign of those effects.

Now for their magnitudes, they are increasing first because incumbents wage offers are increasingly related to ability, and decreasingly to schooling. Second, workers' outside option only depends on schooling during both periods, and the extent to which it depends on schooling remains the same. So on one side, we have incumbent wage offers that increasingly depend on ability and decreasingly on schooling, and on the other side, outside options that steadily depend on schooling. Thus, the difference between incumbents' offers and outside offers increases with ability, and more so over time, while it decreases with schooling, and more so over time. Times 1 and 2 effects of ability and schooling decrease with the uncertainty surrounding the workplace evaluation captured by θ. Indeed, we saw that the wage offers of incumbents are adjusted with a risk premium paid by the worker. Thus, the more variability there is in θ, the lower the optimal wage offer, and the more chances there is that the optimal wage is extended to the worker, because firms are risk neutral profit maximizers. Thus, the more dispersion there is in θ, the more chances there is that workers are selected in the pool of potential quitters; however, those workers that are selected for offers' extension would also have more incentives to quit because the wage offer will be lower. This competing force reduces the independent effect of ability and schooling on quits.

In the next section, I study the second information setup, where the informational gap narrows: at t=2 information is fully symmetric.

Case 2: Decreasing Asymmetry — Information asymmetric at t=1 and fully symmetric at t=2 (ability is observed by outside firms at t=2)

In the same process as for the first case, I start with expanding optimal wages before I discuss quit and layoff rules. I then formulate propositions describing how schooling and ability affect layoff and quit hazard rates through time.

Wage offers. The model is solved with the assumption that outside firms do not observe the signal of ability during the first period of employment but do observe ability during the second period of employment. This yields the following optimal wages at time t=2:

$$v_2^*(z,s) = z \tag{1.29}$$

$$w_2^*(z,s) = \frac{z + v_2^* - \bar{\theta}}{2} = z - \frac{\bar{\theta}}{2} \tag{1.30}$$

Outside firms offer spot contracts. Thus, optimal outside wage v_2^* equals expected productivity, conditional on all information available. Here, outside firms observe ability perfectly at t=2.

The wage schedule of incumbent firms is the same as in the setup 1. Similarly, they charge a risk premium to workers: in fact, incumbent firms insure workers against the risk carried by the shock related to job satisfaction. So the optimal wage from incumbent firms is lower than the optimal wage from outside, although the two expect the same output. Yet, it not obvious that workers will always reject the incumbent wage offer. In fact, the disutility shock can also be a utility shock in the sense that the workplace evaluation is so good that the worker is better off with the incumbent firm. This is also a consequence of the myopia of the worker in this model, which is constraint of the focus on first job.[14]

[14]The implicit assumption is that the impact of the disutility shock are felt when the worker makes the choice between staying and moving. In fact, since a new θ will be drawn in the coming period of employment, the worker is not cleared of all risks.

Before turning to optimal wages at t=1, w_1^* and v_1^*, since w_1^* depends on $E(V_2|\tilde{z}, s)$ and v_1^* conditions on what w_1^* could be, I first discuss $E(V_2|\tilde{z}, s)$, the value that incumbent firms expect at t=2.

$$E(V_2|\tilde{z}, s) = \frac{\bar{\theta}}{8} \tag{1.31}$$

Unlike the case of constant asymmetric, the expected value of the worker at t=2 is fully determined by the variance of θ.[15] In fact, at t=2, the only profit incumbent firms can make originates from the risk premium charged to workers. This is so because their informational advantage disappears at t=2, so there is no information rent to expect. The risk premium goes directly as a profit because firms are risk neutral.

Now at t=1, the information is asymmetric between incumbent and outside firms. Only the incumbent observes a signal of ability $\tilde{z}$ and schooling s, while outside firms only observe schooling s. This yields the following optimal wages:

$$v_1^*(z, s, refuses) = E(z|s) - b \tag{1.32}$$

$$w_1^* = \frac{1}{2}(\tilde{y_1} + E(V_2(z)|\tilde{z}, s) + v_1 - \bar{\theta}) = \frac{1}{2}(E(z|\tilde{z}, s) + E(z|s) - b - \frac{7\bar{\theta}}{8}) \tag{1.33}$$

With outside firms offering spot contracts, their optimal wage at t=1, is built similarly in both information contexts, especially because the information they have access to, at t=1, is the same in both contexts. What differs is the deduction made to compensate for the bad signal sent when a worker quits. In fact, this penalty depends on the what outside firms expect the incumbent is willing to offer to the worker. Because the optimal wage from the incumbent firm at t=1 is not the same in the two informational contexts, the penalty is not the same.[16]

Not surprisingly, the optimal wage from the incumbent is lower in the current context of decreasing asymmetry. In fact, since both incumbent and the outside firms have access to the same information at t=2, there is no informational advantage in-

[15] See Appendix .1.4 for proof.
[16] See Appendix .1.4 for proof.

cumbent firms can benefit from, at t=2. Therefore, incumbent firms have less interest to incentivize the worker by offering a high wage at t=1.

Below, I derive the conditions guiding agents' decisions concerning turnovers. The rules are conceptually the same, but since the information structure and optimal wages differs, the associated separation thresholds also differ.

The layoff rule of the firm. Recalling the timing of events, incumbent firms first evaluate the optimal wage to be offered to each potential worker in the coming period of employment. Then, the former compare the expected profit they can achieve with each of these workers; those who qualify to receive an effective wage offer are those that feature the maximum expected profit. The incumbent firm decides to lay off if the worker currently hired does not meet this criteria. The worker is laid off at time t if:

$$E(y_t|I_t^{incumbent}) - w_t^* < 0$$

Compared to the case of increasing asymmetry, the analytical expression of optimal wages obtained in the current informational context is easier to manipulate. I thus provide the precise lay off condition at each time t=1,2.
The worker is laid off at t=1 if:

$$E(z|\tilde{z}, s) - \frac{1}{2}(E(z|\tilde{z}, s) + E(z|s) - b - \frac{7\bar{\theta}}{8}) < 0$$

$$E(z|\tilde{z}, s) < E(z|s) - b - \frac{7\bar{\theta}}{8} \tag{1.34}$$

Workers that are laid off at t=1 are those for whom the signal received by incumbent firms at t=1, $\tilde{z}$, worsens the prior belief captured by $E(z|s)$.

The worker is laid off at t=2 if:

$$z - z + \frac{\bar{\theta}}{2} = \frac{\bar{\theta}}{2} < 0 \tag{1.35}$$

The lay off condition at t=2 can never be satisfied. In fact, we saw that in this model, firms' rents arise from information uncertainties. What this results suggests is that in the long run, when the market figures out ability of everyone, there is accrued competitiveness and firms are better off not breaking their current match. Because workers are risk adverse and firms risk neutral, firms can only make profit on inside workers by providing insurance to them against the hazardous events that affect workers' appreciation of the workplace.

I now move to the rule concerning separations initiated by workers.

The quit rule of the worker. The worker compares the return from pursuing with the same employer to the return from taking the outside offer. The worker decides to quit at t if:

$$w_t^*(I_t^{incumbent}) - \theta_t < v_t^*(I_t^{outside}) \tag{1.36}$$

The worker quits at t=1 if:

$$w_1^* - \theta_1 < v_1^*$$

$$\theta_1 > \frac{1}{2}(E(z|\tilde{z}, s) - E(z|s) + b - \frac{7\bar{\theta}}{8}) \tag{1.37}$$

The worker quits at t=2 if:

$$w_2^* - \theta_2 < v_2^*$$

$$\theta_2 > -\frac{\bar{\theta}}{2} \tag{1.38}$$

Generally speaking, quits become more and more beneficial as the disutility shock rises. At t=1, information is asymmetric and the critical threshold is specific to each worker. At t=2, information is fully symmetric and the critical threshold is the same for all workers.
To derive the quit and layoff hazard rates at t=2, we need to derive the survival condition at t=1.

The survival condition of the worker at t=1. The worker survives in the job past time t=1 if neither a layoff nor a quit are profitable at time t=1. The worker survives if:

$$E(z|\tilde{z},s) \geq E(z|s) - b - \frac{7\bar{\theta}}{8} \;\;\&\;\; \theta_1 \leq \frac{1}{2}(E(z|\tilde{z},s) - E(z|s) + b - \frac{7\bar{\theta}}{8}) \quad (1.39)$$

Workers that survive are those with good performance and workplace evaluations at t=1. All is now set to derive layoff and quit hazard rates.

Layoff hazard rates. At t=1, the layoff hazard rate is equal to the layoff probability because all workers are in the pool of survivors:

$$\lambda_L(1) = P[E(z|\tilde{z},s) \leq E(z|s) - b - \frac{7\bar{\theta}}{8}] \quad (1.40)$$

There is no layoff at t=2 so the layoff hazard rate is null at t=2.

Quit hazard rates. At t=1, the quit hazard rate is equal to the probability of a quit conditional on no layoff at t=1:

$$\lambda_Q(1) = P[\theta_1 \geq \frac{1}{2}(E(z|\tilde{z},s) - E(z|s) + b - \frac{7\bar{\theta}}{8})|E(z|\tilde{z},s) \geq E(z|s) - b - \frac{7\bar{\theta}}{8}] \quad (1.41)$$

At t=2, the quit hazard rate is equal to the quit probability at t=2, conditional on surviving t=1 and no layoff at t=2:

$$\lambda_Q(2) = P[\theta_2 > -\frac{\bar{\theta}}{2}|E(z|\tilde{z},s) \geq E(z|s) - b - \frac{7\bar{\theta}}{8}, \theta_1 \leq \frac{1}{2}(E(z|\tilde{z},s) - E(z|s) + b - \frac{7\bar{\theta}}{8}] \quad (1.42)$$

Alike the case of constant asymmetry, I now formulate propositions describing how schooling s and ability z affect layoff and quit hazard rates through time. These propositions are hypotheses that will be confronted to the data.

Proposition 3 : *Employer learning and layoffs — Case of decreasing asymmetry*
Holding schooling constant, the layoff hazard rate is negatively related to ability at t=1. The model predicts no layoff at t=2.

Holding ability constant, the layoff hazard rate is positively related to schooling at t=1. The model predicts no layoff at t=2.

Proof 3 *See Appendix .1.4*

In the case of decreasing asymmetry - where incumbent firms have an informational advantage limited to the first period of employment, the predictions on layoffs only identify the effect of schooling and ability at t=1. These predictions at t=1 are qualitatively similar to those obtained in the case of constant asymmetry: layoffs are negatively related to ability and positively related to schooling. Indeed, at t=1, the information structure is the same for the two setups. These two information contexts differ at t=2: in the case one of increasing asymmetry, the informational advantage of incumbent firms accumulates while it disappears in the case of decreasing asymmetry. The assumption of decreasing asymmetry yields no layoffs at t=2, when information is fully symmetric. This outcome of the model puts more emphasis on the fact that layoffs result from information frictions between firms, suggesting that reducing these frictions will also reduce the risk of layoff that workers face.[17] How about quits?

Proposition 4 *: Employer learning and Quits — Case of decreasing asymmetry*
Holding schooling constant, the quit hazard rate is negatively related to ability at t=1. Ability does not affect quits at t=2.

Holding ability constant, the quit hazard rate is positively related to schooling at t=1. Schooling does not affect quits at t=2.

Proof 4 *See Appendix .1.4*

For quits, predictions at t=1 are also aligned with those discussed in the case of constant asymmetry: ability decreases the quit hazard rate while schooling increases it. At t=2, the market has figured out ability of everyone: quits are purely random in the sense they are independent of workers' profile. Thus ability and schooling do not

[17] Because there is no layoffs at time t=2, the effect of ability and schooling is not identified.

matter. Workers' initiated moves are totally governed by the workplace assessment.

In the next section, I implement the empirical test proposed by Farber and Gibbons (1996) to assess the two labor market outcomes of interest: the quit and layoff hazard rates on first job. The objective is to assess how realistic all the predictions discussed earlier are. Besides, the features that are specific to each information setting will be compared to the empirical results to make conjectures on the actual information structure. For the quit hazard rate especially, schooling and ability matter at t=2 in the case 1 (comparable to the short-medium run) ; however, they are irrelevant in the case 2 (approximating the long run).

1.3 Employer learning and First job transition: an empirical test

1.3.1 Empirical strategy

For consistency with the theoretical predictions, I perform a survival analysis and especially focus on estimating conditional hazard rates. Because I want to assess the effect of workers' characteristics on exit hazard rates, I consider a semi parametric proportional hazard function. The targeted labor outcomes relate to job separation on the first job; so the underlying duration is the first job spell.

The test proposed by Farber and Gibbons (1996) consists in estimating two functional expressions of the labor market targeted: one for which the effect of ability is constrained to be constant throughout experience, and another one for which the effect is allowed to change with experience. By comparing the results from these two functional forms, we can assess the trend in the effects of both the hidden skill, which is ability z in our case, and the prior predictor of this hidden skill, schooling s in our case. With t, specifying the underlying duration process, I work with the following hazard functional forms:

$$\lambda^1(t|x(t)) = \lambda_0(t)exp(\beta_{1t}s + \gamma_1 z)$$
$$\lambda^2(t|x(t)) = \lambda_0(t)exp(\beta_{2t}s + \gamma_{2t} z)$$

$\lambda_0(t)$ is the baseline hazard, which only depends on the duration of the spell. It captures the common trend in the hazard rates. For instance, all workers might be less likely to quit or be laid off if they last longer in job, no matter what their skills are. With the multiplicative power, worker's characteristics make the hazard rate rotate above or below this baseline.

Thus, the first functional form is the one for which the effect of ability z is constrained to be fixed throughout experience; γ_1 gives the effect of ability averaged over the job tenure. The second functional form allows the effect of ability z to change with tenure; γ_{2t} measures the effect of ability at tenure t. I present the results of the empirical estimations in this order.

Matching the implications derived from the theoretical model to the empirical model implies:

1. For the case 1: increasing asymmetry
   (a) Layoff: $\beta_{jt} > 0$, $|\beta_{jt}|$ increases with t, $\gamma_1 < 0$, $\gamma_{2t} < 0$, $|\gamma_{2t}|$ increases with t.
   (b) Quit: $\beta_{jt} > 0$, $|\beta_{jt}|$ increases with t, $\gamma_1 < 0$, $\gamma_{2t} < 0$, $|\gamma_{2t}|$ increases with t.

   The parameter γ_1 captures the average effect of ability, while γ_{2t} measures the effect of ability for workers with tenure t. When ability increases while schooling is fixed, the layoff and quit hazard rates decreases. As schooling increases while ability is fixed, the layoff and quit hazard rates increases. Moreover, these effects become larger over time. Since ability becomes more important over time, the model implies that the average effect exceeds the initial one t=1 ($|\gamma_1| = |\gamma_{21}|$).

2. For the case 2: decreasing asymmetry

   (a) Layoff: $\beta_{jt} > 0$, not identified at some point, $\gamma_1 < 0$, $\gamma_{2t} < 0$, not identified at some point.

   (b) Quit: $\beta_{2t} > 0$, $\gamma_{2t} < 0$, $|\gamma_{2t}|$ and $|\beta_{2t}|$ decrease towards 0.

The predictions at t=1 are the same in the two informational settings for both the quit and layoff hazard rates. However, the effect of ability and schooling on the layoff hazard rate, at t=2, is not identified. Besides, ability and schooling are irrelevant to predict the quit hazard rate at t=2; this is workable specificity I will use to tell these two information contexts appart.

1.3.2 Data and summary statistics

The analysis is based on the 1979 cohort of the National Longitudinal Survey of Youth, which is a panel survey of men and women followed from 1979 to 2014. The respondents were 14 to 22 when first interviewed. Relevant information on job history and demographic variables is provided. I restrict the analysis to the random and nationally representative sample. The main information required to test the theoretical predictions is workers' schooling (s), a measure of ability (z), the first unemployment spell, the first job spell, and the reason why the first job has ended.

Each survey-year, respondents provide their highest completed grade of education. Schooling (s in the model) is the highest grade completed that was recorded during the last survey round (2014).[18] Following the literature, I consider individuals with at least 8 years of educational attainment.The reason is that the labor market patterns of workers with less that 8 years of education is generally different.

NLSY79 contains information on AFQT (a measure of ability) scores, which is similar to an IQ test. Following the literature on employer learning, AFQT score is used as a covariate of productivity hidden to the market (z in the model).[19] The AFQT

[18] NLSY79 surveyed respondents annually between 1979 and 1992, and biannually after 1992.

[19] In Lange(2007), some arguments are provided to defend that AFQT is overall not observed by employers. The main idea is that turnover is high for workers with low experience. So performing cognitive skill tests to applicants has poor value.

was administered to the whole sample in 1981. So respondents took it at different ages. To account for potential effects of age and prior schooling, I standardize the scores based on the profile of each age group.

To measure first unemployment spell, I need graduation time, and first job start. I first compute graduation year, as the year in which the respondent graduated from highest level of schooling ever recorded. I assume the graduation happened in June. For each job, defined as an employer-employee match, the year, month and day of start and end are recorded; I only used month and year to have time in months. First unemployment spell is measured as the number of months since the graduation time to the start of first job start; first job is the job that started the earliest after graduation.[20]

For each job, the NLSY also asks the respondents the reason for separation. I label the job end reason as a layoff when the reasons given are: "Layoff, job eliminated", "Layoff", "Discharged or fired". I label the job end reason as a quit when respondents answer "Quit because found a better job", "Quit to take another job", "Quit to look for another job", "Quit because wages too low", "Quit because of employment conditions (didn't like work, hours,conditions, or location,etc)". [21] To compute the duration process for which layoff is the failure event, I assumed that a job that terminated for a motive different from a layoff is a censored observation. I did the same for quit.

NLSY79 also records occupation and industry on a survey-year basis. I consider the industry and occupation reported during the last interview before the job ended.[22] These variables are used as controls along with race, gender and macroeconomic conditions captured by the dummies for the year the job ended.[23]

[20]Jobs held as students are overlooked since they happen before the worker left school for good.

[21]There is in total 41 categories of reason in the data. Some examples are "Plant closed", "End of temporary or seasonal job", "Quit for pregnancy or family reasons", "Quit because of respondent's ill health, disability, or medical problems".

[22]There is the concern that workers' sorting into jobs be correlated with both their characteristics and the mobility rate associated with specific jobs; I therefore add occupation and industry fixed effects for all the three outcomes as a way to control for it. However, a full analysis of such sorting is out of the scope of the current study.

[23]The survey has zipcode level data, but it's confidential and restricted access. I don't have restricted access to the data on geographic location. Yet, controlling for rural/urban area and unemployment rates at the local level would have been interesting to capture differences in the labor market conditions that the workers face.

Table 1.1 summarizes the main variables used. The average years of schooling in the sample is 13.4, but 50% of the sample did not graduated from high school (median=12), and dropped out between 8 and 12 years of education. The AFQT scores were standardized to have mean 0 and standard deviation 1.[24] The mean first unemployment spell is 11 months, and the mean first job spell is approximately 47 months.[25] There is high dispersion in the first job spell: the standard deviation is equal to 67 months. Comparing standard deviations, it does not seem like differences in schooling and ability could explain the high variation in first job spell. Looking at the covariance between ability and schooling, the hypothesis that schooling is a fair predictor of ability makes sense. Layoffs represent 13% of first job exit motives, and quits 12% out of other motives considered in NLSY data, such as health purposes. The sample is balanced with respect to gender but not with race; estimations are made on the random sample of NLSY79.

Table 1.1: Summary statistics

Variable	Mean	Std. Dev.	Median	Min.	Max.	N
s	13.4	2.37	12	8	20	5282
z	0	0.99	-0.02	-1.93	2.2	4999
Unemp. spell	11.03	7.65	11	0	42	5282
Job spell (1st)	46.80	67.16	22	1	487	5282
Cov(s,z)						0.5961
1st job end motive						Layoffs 13% ; Quits 12%
Gender						Female 50% ; Male 50%
Race						White 87.51% ; Black 12.49%

Notes. Descriptive statistics computed by the author on the sample used to perform most estimates: the subset of NLSY-79 random sample with exclusions due to missing values and minimum schooling.

[24]The standard deviation is not exactly 1 in the table because the standardization was completed before restricting the sample to people with at least 8 years of education.

[25]Although the model does not cover job finding, I also perform the test proposed by Farber and Gibbons (1996) on the first job finding hazard rate. The underlying duration is the first unemployment spell. When applied to first job finding, the exercise is actually testing whether ability is relevant (or observed, it is not possible to tell appart) in assessing workers' chances to find a first job.

1.3.3 Results

In this section, I explore the empirical patterns of layoffs and quits to gauge the match between them and the theoretical predictions summarized earlier. Moreover, the results of estimations will help make conjectures about the actual informational asymmetries that the cohort of workers under study have been confronted to. The model under study does not deliver workable predictions on job findings because hiring is granted to all workers in the model. However, I also perform the same test on job finding and proceed reversely: I discuss the empirical results first and then sketch model inputs that could deliver the empirical facts identified in between.

1.3.4 Layoffs

Table 1.2 shows the estimated effects of schooling and ability on the layoff hazard rate. These results are also presented on the two figures 1.2 and 1.3. The sample is restricted to workers for which the first job lasted at most 16 years. The dependent variable is the layoff hazard rate from the first job spell. The coefficients are estimated by Maximum Likelihood. Columns differ by control sets with the basic set made of gender, race and macroeconomic conditions. The columns 1 and 2 report the specifications with the basic set of controls.[26] [27] Referring to the empirical strategy outlined in the section 1.3.1, columns 1, 3 and 5 report the estimation of the hazard rate λ^1, while column 2, 4 and 6 report the estimation of the hazard rate function λ^2. The interaction terms with job tenure are splitted in two-year bins.

From column 1, we see that, for the first two years of tenure, one additional year of schooling reduces the layoff hazard rate by 22.5 percent, with a standard deviation of 2.1 percent. This figure drops to 12 percent during the two subsequent years of tenure, the standard deviation remains the same. The estimated effect of schooling on the layoff hazard rate, from 4 to 10 years of job tenure, is statistically insignificant.

[26] Additional controls include occupation and industry fixed effects as a way to control for workers' sorting into jobs.

[27] Controlling for rural/urban area and unemployment rates at the local level would have been interesting to capture differences in the labor market conditions that the workers face. However, information on zipcode is restricted.

Table 1.2: Effect of Schooling and Ability on Layoff hazard rate

	(1)	(2)	(3)	(4)	(5)	(6)
schooling× exp (0-2 yrs)	-22.5 ***	-22.2***	-22.2***	-22.6 ***	-19.8***	-19.2***
	(2.1)	(2.1)	(3.2)	(3.3)	(3.5)	(3.6)
schooling× exp (2-4 yrs)	-12.0**	-13.5**	-9.4	-13.0*	-5.1	-7.8
	(5.8)	(6.5)	(6.4)	(7.4)	(6.2)	(6.9)
schooling× exp (4-6 yrs)	-9.4	-3.5	-10.0	-2.9	-5.2	3.2
	(8.1)	(9.8)	(7.8)	(9.0)	(8.2)	(9.6)
schooling× exp (6-8 yrs)	-8.1	2.9	-8.2	3.4	-3.8	11.2
	(7.4)	(9.5)	(7.3)	(9.3)	(8.3)	(10.2)
schooling× exp (8-10 yrs)	-0.9	7.9	3.3	11.0	5.7	14.5
	(6.8)	(8.2)	(6.8)	(8.9)	(7.4)	(10.1)
schooling× exp (10-12 yrs)	12.9*	11.1	13.3*	6.9	12.6	14.1
	(7.7)	(9.4)	(7.5)	(8.9)	(8.0)	(9.7)
ability	-28.5***		-32.9***		-29.6***	
	(6.4)		(7.8)		(7.7)	
ability× exp (0-2 yrs)		-13.0		-3.0		1.82
		(13.4)		(15.1)		(14.8)
ability× exp (2-4 yrs)		-22.6		-18.0		-19.2
		(14.3)		(15.5)		(15.4)
ability× exp (4-6 yrs)		-45.8***		-52.4**		-53.6***
		(17.4)		(16.9)		(17.5)
ability× exp (6-8 yrs)		-70.1***		-74.0**		-84.3***
		(22.8)		(23.3)		(24.0)
ability× exp (8-10 yrs)		-69.3***		-62.9**		-66.1**
		(26.3)		(29.3)		(31.7)
ability× exp (10-12 yrs)		-21.1		-2.9		-88.1**
		(28.2)		(35.1)		(53.4)
Year fixed effects	Y	Y	Y	Y	Y	Y
Gender	Y	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y	Y
Occupation	N	N	Y	Y	N	N
Industry	N	N	N	N	Y	Y
Observations	4,300	4,300	3382	3382	3473	3473

Notes. The dependent variable is the layoff hazard rate from first job spell after graduation. The sample is restricted to workers for which the first job lasted at most 16 years, but I only reported coefficients for up to 12 years. The reported parameters and standard deviations are those estimated multiplied by 100. Specifications differ by control sets: Y(es) when the control variable is included, N(o) otherwise. The main explanatory variables are ability measured by AFQT scores and schooling measured by years of education, both interacted with job tenure (in years) dummies. The set of observations is smaller when job controls (industry and occupation) are included because of coding changes throughout NLSY time series; so for specifications (2), (3) and (4), I have restricted the sample to jobs held before the coding switch. *,**, *** respectively refers to 10%, 5% and 1% significance level.

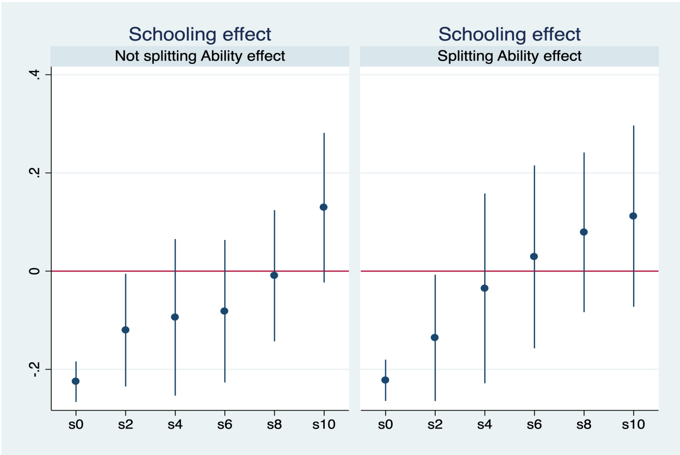


Figure 1.2: Effect of schooling on layoff hazard over job spell

Notes. The figure shows the marginal effect of schooling on the layoff hazard rate by years of tenure. The left panel gives the effect of schooling, when the effect of ability is fixed over first job tenure, Table 1.2, column 1. The right panel does the same but with the effect allowed to vary with job tenure, Table 1.2, column 2.

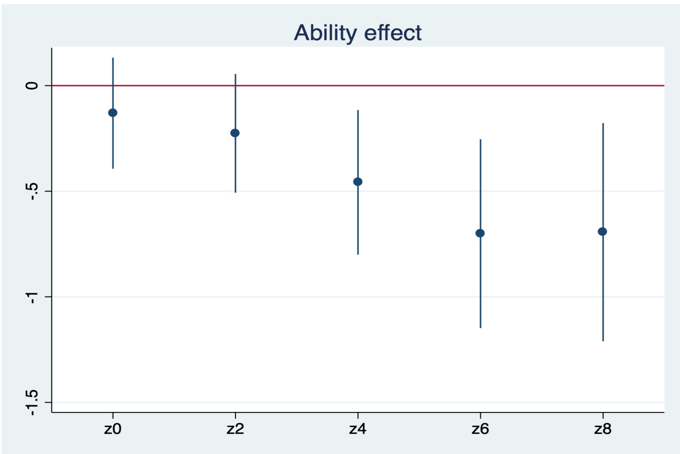


Figure 1.3: Effect of ability on layoff hazard over job spell

Notes. The figure shows the marginal effect of ability on the layoff hazard rate by years of tenure, see Table 1.2, column 2 for values.

It is evaluated at 9.4 percent for 4 to 6 years of tenure, at 8.1 percent for 6 to 8 years of tenure, and at 0.9 percent for 8 to 10 years of tenure. The estimated effect of schooling from 10 to 12 years of tenure, is large, and quite precisely estimated: one additional year of schooling increases layoff hazard rate by 13 percent. The sign and the large magnitude coincide with the predictions of the model for the case of increasing asymmetry. In fact, the case of increasing asymmetry portrays the state of the labor market in the short-medium term, when asymmetry is pronounced and incumbents have a strong informational advantage. Over tenure, incumbents become more informed of their workers' actual skills. Thus, for two workers with identical skills, the one with higher schooling has a higher market value. To keep such a worker, the incumbent has to pay a higher wage which results in a lower profit; such a worker faces a relatively higher layoff risk.[28]

The estimates for schooling are closer to the theoretical predictions of the model with increasing asymmetry. This is so especially for those concerning the change in the effect of schooling: $\beta_1, \beta_2 > 0$. The estimate of the initial effect has a reversed sign from the one predicted by both versions of the model. The empirical results imply that for two workers with same ability, the one with higher schooling faces a lower risk to be laid off in the first years tenure. Thinking through the model, such a result could arise if for instance initial signals of ability ($\tilde{z}$) are so noisy that schooling takes more importance in firms expectations', for the first years of tenure. At some point, for tenures above 10 years, the sign of the estimate switches and retrieve the one predicted in the model. One plausible explanation of the positive effect of ability is over-education. Indeed, since the specification controls for ability, increasing schooling above a certain threshold implies that the prior belief of the firm was an overestimation of the ability of the worker. For such a worker, as the job lasts longer, the incumbent firm corrects its initial belief, putting the worker at a higher risk to be laid off.

As for ability, we see that, irrespective of the job duration, one standard deviation increase in ability reduces the layoff hazard rate by 28.5 percent. Interpreting this result requires to discuss the results of column 2.

[28] Actual pay schedules incorporate a premium for each education grade. The model is consistent with this fact.

Column 2 reports the effect of schooling and ability when the effect of ability is allowed to change over the job tenure. The results on the effect of schooling are overall similar to those obtained with the first specification. However, the magnitude of the effect decreases faster after 4 years of job tenure. Also, the switch in the sign of the coefficient happens earlier: for tenures longer than 6 years, schooling is positively related to layoffs. However, for the results displayed in column 1, the switch in sign happens after 10 years of job tenure. This indicates that the negative effect of schooling estimated in the first specification (column 1) for tenures between 6 and 10 years was spurious. This makes sense because schooling is correlated with ability: when the effect of ability is assumed to be fixed while the one of schooling is allowed to change, schooling picks some of time dependent effect of ability.

The estimated effect of ability on the layoff hazard rate from 0 to 4 years of tenure is statistically indistinguishable from 0. We can conclude that the effect is near null because for half of the sample, first job lasted less than 22 months, so less than 2 years. This means that there is enough data to estimate the parameter precisely. However, one standard deviation increase in ability reduces the layoff hazard rate by 45.8 percent, for tenure ranging between 4 and 6 years. This figure jumps to 70.1 percent for tenure between 6 and 8 years, and decreases slightly to 69.3 percent, for tenure between 8 and 10 years. For tenures between 10 and 12 years, the coefficient is estimated at 21.1 with a high standard deviation of 28.2 percent. Overall, these results on the effect of ability are also closer to the assumption of increasing asymmetry. Especially, the average effect of ability is larger that the initial effects ($|\gamma_1| > |\gamma_2|$). In fact, in the long run, if the market figures out ability of everyone (case of decreasing asymmetry), we should expect less layoffs over time. The effect of ability is estimated with such precision that it is unlikely that the long run version of the model be the data generating process.

Taken together with the results of column 1, the pattern of the effect of ability is aligned with the overall predictions of employer learning. First, the averaged effect of ability, estimated in column 1 is larger than the effect of ability during the first years of job tenure: this effect is even most likely null. Thus, ability is only a marginal predictor of layoff during the first four years of job tenure. It does explain a big part of the lay off risk from 4 to 10 years of tenure. This is consistent with the idea that firms value workers' skills in the way they assess workers. Yet, they do not observe them initially; they only learn them through experience.

Columns 3 to 6 add industry and occupation controls to the specifications 1 and 2. The results are qualitatively the same. Looking at the results reported in column 3 or 5 and comparing them with column 1, we note that the average effect of ability on the layoff rate was underestimated in absolute value, when occupation or industry respectively, was not controlled for. Also worth discussing is the specification 6 which includes industry controls and allows the effect of ability to change over time. We see that the decrease in the effect of schooling as well as the sign switch, happen faster. In addition, as job tenure rises, the increase in the predictive power of ability is stronger. Besides, the effect of ability for tenures between 10 and 12 years, is precisely estimated at a 88.1 percent decrease in the layoff hazard rate as the result of a one standard deviation increase in ability.

It is not surprising that industry and occupation are confounding factors. In fact, these results are consistent with education and ability being correlated with industry and occupation, but also with layoffs happening relatively more for some occupations, and in some industries.

Overall, updates to employers' beliefs about workers seem to matter in their decision to lay workers off. What the different specifications we discussed above suggest is that initial signals of ability are so noisy that employers' beliefs are driven by schooling when tenure is still short. As the employer and the worker spend time collaborating, the expectations of the employers over the productivity of the worker correlates less with schooling because hidden skills become more and more perceptible and determinant to decide on the value to pursue the job match. At the same time, because learning is asymmetric, workers with high schooling have high market value, forcing incumbents to pay them high wages. Besides, these results imply that the market is still in the short-medium run, and information frictions between incumbents and outside firms are still present. Still, if the labor market is not competitive, or if the production includes a match specific component, it is possible to have such results even if information is fully symmetric. Indeed, the predictions discussed earlier are true in the case of identical production functions and the labor market competitiveness.

1.3.5 Quits

Table 1.3 reproduces the table 1.2 for the quit hazard rate. It displays estimated effects of schooling and ability on the quit hazard rate. These results are also presented on a figure format, figures 1.4 and 1.5. The sample is restricted to workers for which the first job lasted at most 16 years. The dependent variable is the quit hazard rate from first job spell. As before, the coefficients are estimated by Maximum Likelihood. Columns differ by control sets with the basic set made of gender, race and macroeconomic conditions. Here, columns 1 and 2 report the specifications with the basic set of controls.[29] Referring to the empirical strategy outlined in the section 1.3.1, column 1, 3 and 5 report the estimation of the hazard rate λ^1, while column 2, 4 and 6 report the estimation of the hazard rate function λ^2. The interaction terms with job tenure are splitted in two-year bins.

We see from column 1 that the propensity to quit on the first job is positively related to schooling. More specifically, one additional year of education increases the quit hazard rate by 30.2 percent, for the first two years of tenure. This statistic was estimated with a standard deviation of 3.9 percent. The effect of schooling remains positive, but decreases monotonically over the job spell; one additional year of schooling increases the quit rate by 13.6 percent for tenures between 8 and 10 years.

These patterns are partially aligned with the theoretical predictions from the two versions of the model; the case of decreasing asymmetry matches the predictions closer. Indeed, both versions imply that schooling has a positive effect on the quit hazard rate. When the asymmetry widens this effect increases with job spell. This is the consequence of the learning asymmetry: as the job lasts longer, workers' outside option responds more to schooling than their current compensation do. In fact, since the incumbent firm learns faster, schooling weights less and less in the expectations. When the asymmetry narrows, this effect decreases with job spell, and schooling

[29] As for layoffs, I add occupation and industry fixed effects to control for the sorting of workers into jobs.

Table 1.3: Effect of Schooling and Ability on Quit hazard rate

	(1)	(2)	(3)	(4)	(5)	(6)
schooling× exp (0-2 yrs)	30.2***	32.9***	17.2***	20.2***	27.7***	30.0***
	(3.9)	(4.7)	(4.9)	(5.6)	(4.8)	(5.4)
schooling× exp (2-4 yrs)	26.8***	30.6***	14.6***	16.8***	23.4***	24.3***
	(4.4)	(4.4)	(5.2)	(5.7)	(5.1)	(5.5)
schooling× exp (4-6 yrs)	27.1***	29.8***	14.0**	15.1**	22.1***	21.1***
	(5.2)	(6.5)	(5.5)	(6.5)	(6.1)	(6.9)
schooling× exp (6-8 yrs)	18.6***	11.3	7.8	-6.2	19.4***	13.8
	(5.4)	(7.3)	(5.6)	(7.5)	(6.7)	(9.9)
schooling× exp (8-10yrs)	13.6**	10.3	13.0*	16.1*	12.9*	2.2
	(6.3)	(7.2)	(7.4)	(9.0)	(7.4)	(7.8)
schooling× exp (10-12 yrs)	4.0	1.9	2.9	2.1	6.8*	5.6
	(2.8)	(3.1)	(3.4)	(3.5)	(3.9)	(3.9)
ability	-4.4		-7.9		-2.8	
	(6.2)		(7.9)		(7.7)	
ability× exp (0-2 yrs)		-15.7		-20.1		-12.2
		(11.5)		(13.7)		(13.6)
ability× exp (2-4 yrs)		-22.3		-15.7		-4.8
		(17.2)		(19.1)		(19.2)
ability× exp (4-6 yrs)		-15.2		- 11.1		2.1
		(21.1)		(19.9)		(19.1)
ability× exp (6-8 yrs)		26.9		48.6**		25.1
		(22.3)		(24.2)		(25.5)
ability× exp (8-10 yrs)		7.9		- 21.5		37.8
		(20.1)		(18.3)		(27.6)
ability× exp (10-12 yrs)		21.2		19.8		6.7
		(21.0)		(24.8)		(25.3)
Year fixed effects	Y	Y	Y	Y	Y	Y
Gender	Y	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y	Y
Occupation	N	N	Y	Y	N	N
Industry	N	N	N	N	Y	Y
Observations	4,999	4,999	3,802	3,802	3,915	3,915

Notes. The dependent variable is the quit hazard rate from first job spell after graduation. The sample is restricted to workers for which the first job lasted at most 16 years, but I only reported coefficients for up to 12 years. The reported parameters and standard deviations are those estimated multiplied by 100. Specifications differ by control sets: Y(es) when the control variable is included, N(o) otherwise. The main explanatory variables are ability measured by AFQT scores and schooling measured by years of education, both interacted with job tenure (in years) dummies. The set of observations is smaller when job controls (industry and occupation) are included because of coding changes throughout NLSY time series; so for specifications (2), (3) and (4), I have restricted the sample to jobs held before the coding switch. *,**, *** respectively refers to 10%, 5% and 1% significance level.

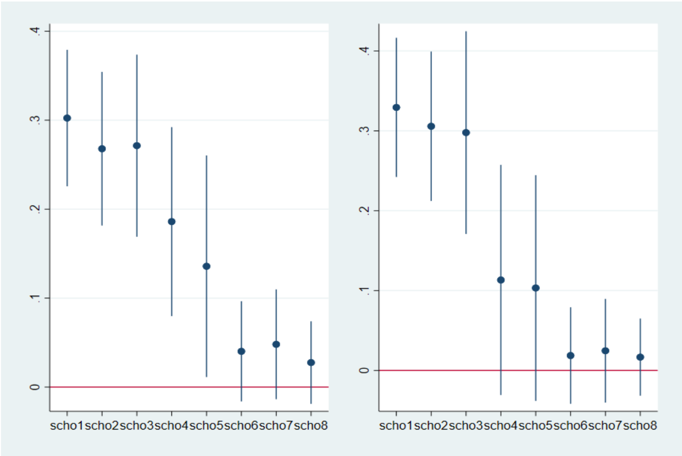


Figure 1.4: Schooling effect on Quit hazard over job spell

Notes. The figure shows the marginal effect of schooling on the quit hazard rate by years of tenure. The left panel gives the effect of schooling, when the effect of ability is fixed over the first job tenure, see Table 1.3(column 1) for values. The right panel does the same but with the effect allowed to vary with job tenure; see Table 1.3, column 2 for values.

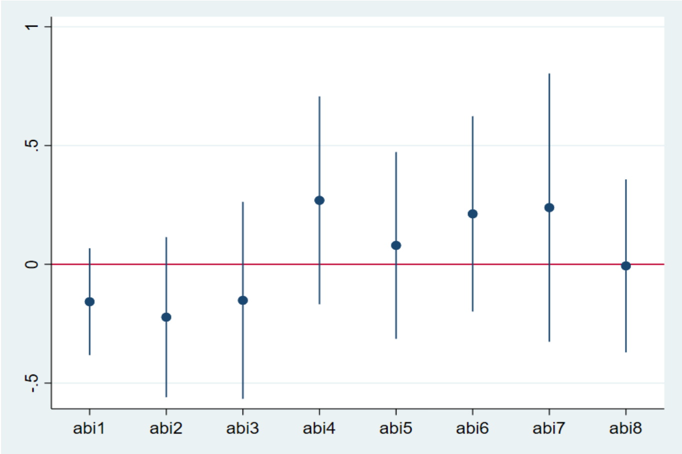


Figure 1.5: Ability effect on Quit hazard over job spell

Notes. The figure shows the marginal effect of ability on the quit hazard rate over years of tenure; see Table 1.3, column 2 for values.

is irrelevant at the limit. This is the consequence of full learning: incumbent and outside firms have access to the same information at the limit, so that outside and inside option respond identically to schooling. Thus, quits are purely random. Based on what the data suggests, newly hired workers tend to quit more when they have high education. But, as the job lasts longer, schooling does not matter that much in their quit decision. This result weights in favor of the hypothesis of decreasing asymmetry. These results are also consistent with human capital mechanisms such as in Kaymak(2004). In his model, workers invest in post-schooling training to compensate for the lack of schooling. If formal schooling and training are substitutes, then the relevance of schooling for labor market opportunities would fade away. So we should see schooling predictive power decrease, as it is the case here. This model also requires a production function that is not fixed over time. In such a case, symmetric information will be compatible with both the results on quits and the results on layoffs.

Turning to the effect of ability from specification 1, we see that, irrespective of the duration of the job, a one standard deviation increase in ability reduces the quit hazard rate by 4,4 percent, estimated with a standard error of 6.2 percent. Because we have enough observations, we can conclude that overall, ability does not matter for quit. This result alone is not enough to assess the learning hypothesis and the information asymmetries.

In the specification reported in column 2, ability is interacted with job duration to evaluate how its effect changes every two years. The results on the effect of schooling are qualitatively the same as those reported in column 1; it appears more clearly on the figure format. For tenures between 0 and 6 years, the estimates have a larger magnitude. The effect of schooling on quit is less precisely estimated for tenures larger than 6 years.

Splitting the effect of ability biyearly does not add much to the results of column 1. Recalling that half of the sample has job tenure below 2 years, we can conclude that ability does not matter during the first two years of tenure. Beyond two years, it is difficult to conclude because the statistical insignificance can result from the fact that ability does not matter, or from the fact that there are fewer observations. If we assume the latter, then increasing ability reduces the quit hazard rate for tenures between 2 and 6 years. Above 6 years, ability is positively related to the quit propensity.

Concerning the effect of ability on the quit hazard rate, what stands out from columns 1 and 2 is that ability does not matter much to explain workers' quit behaviour on their first job. However, should ability matter, it reduces the quit rate for up to 6 years of tenure. Once tenure exceeds 6 years, when ability increases, the quit propensity also increases.

The results from columns 1 and 2 are not conclusive: some predictions of the model versions of model are confirmed in the data, others are not. The estimations show that schooling is positively related to quit; this is also the case in the theoretical predictions of both versions of the model. The change in the magnitude gives credit to the hypothesis of decreasing asymmetry. In the data, the effect of schooling decreases; the model with decreasing asymmetry also predicts a such a trend. As for ability, its effect on the first job is not precisely estimated in our sample. We can venture to say the sign of the effect was correctly estimated, in which case it is aligned with the theoretical predictions for workers with tenure below 4 years. For experienced workers, the empirical results are at odds with employer learning. In fact, from the theoretical predictions, the effect of ability is negative. This effect becomes larger with the job spell in the case of increasing asymmetry. This is so because workers with high ability see their compensation increase over time as incumbents learn, while their outside option remains the same because outside firms do not learn. In the case of decreasing asymmetry, the effect of ability tends to zero because incumbent and outside firms observe ability at the limit, so that quits are purely random.

The results are not significantly modified when occupation and industry controls are added-columns 3 to 6. Worth noting is that when occupation is controlled for, the estimate of the effect of ability on the quit rate for workers with 6 to 8 years of tenure is positive, high, and statistically significant.

These results altogether suggest that the quit behaviour of workers with limited experience is influenced by the learning process of employers, which converges to a situation where the market has figured out ability of everyone. These patterns are also consistent with a situation whether job search effort matters, and workers have preferences for occupations. In fact, one can imagine that workers with high skill put less effort in job search early on but more as the unemployment spell lasts longer. The fact that ability decreases the quit rate initially but increases it later on supports this mechanism because it nicely fits into it. Indeed, workers with high skill that might have experienced longer unemployment spell, could be reluctant to quit in their first years of experience. Once they have acquired experience, they are more likely to quit to return to their preferred occupation.

1.3.6 Job Finding rates

The estimated effects of schooling and ability on the job finding rate are reported in table 1.4. The dependent variable is the exit hazard rate from first unemployment spell after graduation - exit referring to starting a paid job. The coefficients are estimated by Maximum Likelihood. Specifications differ by control sets: all specifications include as controls, gender, race and macroeconomic conditions captured by fixed effects of the year the job ended. Specifications in columns 2-4 add indicators for industry and occupation categories.

The column 1 refers to the specification with the basic set of controls: gender, race and macroeconomic conditions captured by fixed effects of the year the job ended. We see that one additional year of schooling raises the job finding rate by 7.5 percent, with a standard error of 2.8 percent. The coefficient on the interaction between schooling and unemployment duration is virtually zero, indicating that the effect of schooling remains stable throughout the unemployment spell.

Since schooling is a correlate of ability, this is consistent with the idea that firms are interested in the expected productivity of the worker when making their hiring decision, but with no additional signal of productivity being observed as the worker remains unemployed.

One standard deviation increase in ability reduces the job finding rate by 10.4 percent, with a standard error of 6.2 percent. The coefficient on the interaction is also zero, confirming that no additional signal is observed as workers remain unemployed.

Table 1.4: Job finding Hazard

	(1)	(2)	(3)	(4)
schooling	7.5***	9.0**	8.2**	9.2**
	(2.8)	(4.5)	(4.1)	(4.6)
schooling × unemp	- 0.6	-.0.1	-1.45	-1.04
	(0.5)	(0.9)	(0.9)	(0.9)
ability	-10.4*	-9.2	-14.7*	-1.44
	(6.2)	(8.6)	(8.4)	(9.1)
ability×unemp	1.2	1.9	3.1	2.6
	(1.2)	(1.9)	(3.1)	(2.6)
Year fixed effects	Y	Y	Y	Y
Gender	Y	Y	Y	Y
Race	Y	Y	Y	Y
Occupation	N	Y	N	Y
Industry	N	N	Y	Y
Observations	4,300	3,382	3,473	2,425

Notes. The dependent variable is the exit hazard rate from first unemployment spell after graduation. The main covariates are schooling, measured by years of education, ability captured by standardized AFQT scores, and their interaction with unemployment duration measured in years. The reported parameters and standard deviations are the estimated multiplied by 100. Columns differ by control sets: Y(es) when the control variable is included, N(o) otherwise. The set of observations decreases when job controls (industry and occupation) are included because of coding changes; so for specifications (2), (3) and (4), I have restricted the sample to jobs held before the coding switch. *,**, *** respectively refers to 10%, 5% and 1% significance level.

The fact that the initial effect of ability is negative, is intriguing. This result seems at odds with employer learning. Besides, the estimated parameter driving the change in the effect of ability is statistically insignificant; but it is positive and relatively large (twice the estimate of the same parameter for schooling). This implies that the negative effect of ability on job finding decreases over unemployment spell. These results taken together with those on schooling, suggest that some unobserved characteristics of workers make them less likely to find a job when they have high ability, holding schooling constant. As raised earlier, they might for instance invest less effort in job search. One reason could be over-confidence. The fact that the negative effect of ability on job finding rate decreases over unemployment spell is consistent with the search effort mechanism. Indeed, high ability workers might ultimately put more effort as they fail finding a job or an occupation early. If workers with high ability put less effort in job search, it could also be because they prioritize self employment once they complete school. Thus, it would be interesting to investigate whether ability predicts the probability that graduates' first best occupation is self-employment.

The results of column 3, which includes industry as a control, are qualitatively similar to those of column 1. However, the predictive power of schooling and ability is stronger. Thus, industry was a confounding factor.

Results from columns 2 and 4 provide interesting insights and reconcile the different mechanisms discussed above. Indeed, once occupation is controlled for, ability becomes statistically insignificant, both at baseline, and throughout the unemployment spell. These patterns are consistent the hypothesis that employers are uncertain about workers' actual skills and rely on schooling to make their hiring decision. It is also consistent with job search effort being correlated with both workers' ability and preferences in occupations.

A better test of employer learning and job finding requires to look at multiple jobs and see whether over time, it becomes easier for workers with high ability to find a job; this will weight in favor of the hypothesis of learning. The test I am proposing is not fully conclusive about whether workers' ability is even something the firms care about; it does suggest that schooling increases the chance to be hired and that this effect remains constant over unemployment duration. If we add the existing results on wages to the picture, we can conclude that firms do value ability. Indeed, wages are increasingly related to ability over experience. If workers are compensated based on their cognitive skill, we can plausibly conclude that firms value these skills. Therefore, the results confirm that first job finding is subjected to uncertainty about workers' productive skills, and that no learning happens throughout the unemployment spell.

1.4 Conclusion

I examined the implications of employer learning for job turnover in a setting where incumbents have an informational advantage over outside employers. I consider two cases: one where this informational advantage widens over time, and another one where it narrows and disappears at the limit. The empirical patterns of quit and layoff rates across workers with different schooling and ability levels are consistent with asymmetric learning: mobility appears to be relatively higher among educated workers compared to workers at the higher tail of the productivity distribution. Some features of the data suggest that the asymmetry intensifies over time. The version with decreasing asymmetry matches the data only if early signals of ability are very noisy or production processes are not the same across firms.

In particular, ability has a negative effect on layoffs; it becomes a stronger predictor over time. In addition, the sign and the trend of the effect of schooling on layoffs match the predictions for experienced workers; however the estimates are not precise. Some patterns on quit rates are also supportive of the model, and especially the case of decreasing asymmetry. Indeed, the quit rates depend positively on schooling and negatively on ability among newly hired workers, consistent with both versions of the model. Schooling is less relevant to predict the quit behavior of senior workers: this is in accordance with random quits that would happen in the long run when the information frictions disappear and the market has figured out ability of everyone. Ability is also less relevant to predict the quit behavior of senior workers, but not only the estimates are imprecise but the sign is opposite to the predictions. These specific results could be reconciled by incorporating workers' job search effort. The idea of job search effort is also substantiated by the results on job finding.

This research aimed to complete the picture describing how workers' outcomes would be affected if employers don't observe their skills as they enter the labor market. I investigated the impact of updates to firms' beliefs on workers' first job finding, quit and layoff probabilities. Previous studies focused on the evolution of wages over experience. But job transition is important especially first job finding and layoff risk, because they shape subsequent job opportunities. Using a model in three periods, I addressed the question in a context of asymmetric learning (incumbent and outside firms learning at different pace), with schooling as a predictor of skills available to all protagonists. I studied two versions of the model. In the first one, the asymmetry in information deepens. In such a context, only the incumbent firm learns, making workers with high schooling always willing to quit, to get a reset. In the second one, the asymmetry in information shrinks and disappears at the limit, so that turnover is purely random. In the empirical part, I perform a survival analysis to test the predictions of the model using the 1979 cohort of National Longitudinal Survey of Youth.

One important message of this study is that highly productive workers may incur layoffs during their first years of experience because of erroneous assessments of their performance. Such noisy assessments are very likely for newly hired workers. Based on the findings, a policy recommendation to reduce the adverse effects of such information frictions would be to subsidize employment trials to help workers go through the first years of their working lives. Moreover, rotating task assignments between for instance team and individual work will give more precise signals of workers' abilities and speed up the learning process.

Chapter 2

Immigration Shock and Sustainable Debt[1]

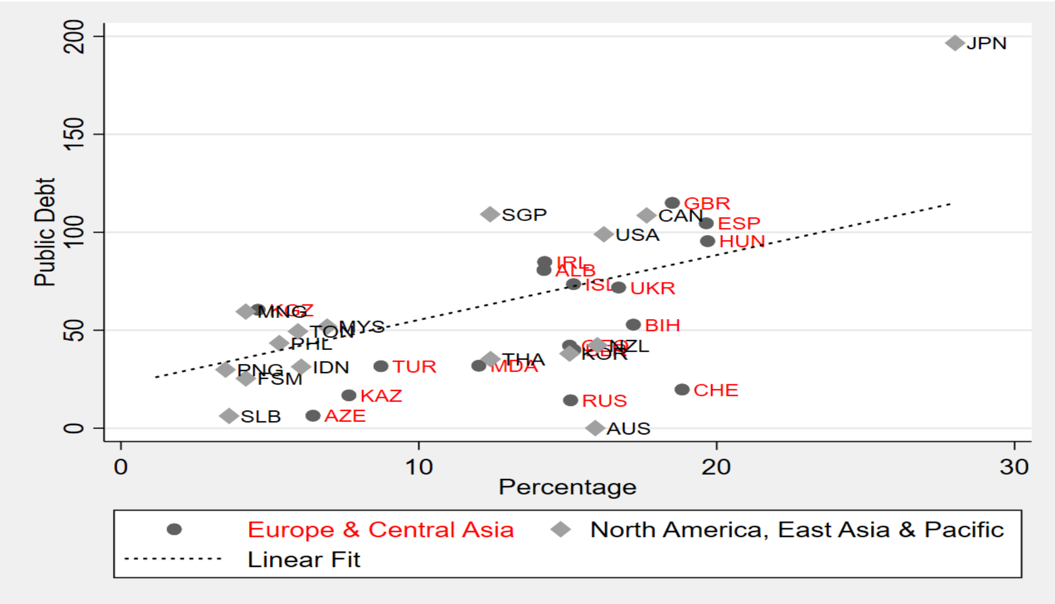
2.1 Introduction

Evidence of a relationship between population age profile and public debt burden in OECD countries has become increasingly apparent over the past two decades. Indeed, during this same period, the population of these countries began to age: not only has the median age of Western society increased by more than 5 years, but also the proportion of older people (who are often large recipients of public benefits) has become increasingly important. Based on World Bank Data Catalog, there appears to be a positive relationship between the level of public debt and the proportion of the population aged 65 and over.

As the scientific society became aware of the link between tax burden and demographic change, it became important to rigorously examine reform proposals that may directly or indirectly affect debt sustainability. Immigration reforms are an example of these, as they can change expectations in terms of fiscal balances. This issue has received increasing attention, in both public and scientific debates. In Denmark

[1]This chapter is a co-authored work with Guy Arnold Djolaud (McGill).

Figure 2.1: Public Debt as % of GDP vs 65 + Population as % of Total Population, most recent data



Sources: From Authors based on World Bank Data Catalog.

2001 elections, immigration policy was a key topic; an argument to limit immigration was the resulting tax burden. Previous studies find that the fiscal contribution of the immigrant population as a whole is quite small (Rowthorn (2008)). However, once age and education are taken into account, young and highly skilled immigrants generate significant net contributions, while low-skilled retirees give rise to significant costs (Lee and Miller (2000) and Storesletten (2000)). Indeed, as might be expected, when immigrants enter their working lives, they make a net contribution to retirees through tax payments. But these immigrants will eventually retire and receive pensions, the present value of which might or might not outweight their positive contribution during the work period. This raises a concern about the long-run sustainability of policies that rely on skilled immigration to close short-to-medium-run fiscal deficits.

Most studies dealing with the macroeconomic effects of immigration do not directly address sustainable debt, as defined by D'Erasmo et al. (2016)-put simply, sustainable debt is that initial level of debt that is covered by government present discounted value of all primary balances. Moreover, they focus on the positive aspects of immigration. The contribution of our study is therefore diverse. First of all, it is original in the sense that it addresses the question of debt sustainability, following an immigration shock. In addition, it proposes an elaborate theoretical model. Indeed, previous immigration studies only include heterogeneity in the production of agents (skilled or unskilled) and do not take into account the fact that the age structure of immigrants have different macroeconomic consequences, especially when looking at the effects on public finances through social security. So we propose a Dynamic Stochastic General Equilibrium model with period of inactivity, working period and retirement. We provide a theoretical assessment of the implications of immigration on sustainable debt and asset prices. Following Conesa and Garriga (2008), we incorporate efficiency of work that decreases with age. We also consider that work efficiency differs by immigration status. The study by Krieger (2004) showed the importance of considering fertility when studying the macroeconomic effects of immigration; we moreover assume that immigrants fertility is higher. Our model is a modified version of D'Erasmo et al. (2016) study on sustainable debt, with different dimensions of agents' heterogeneity.

Using Canada aggregate data on population structure and skill distribution, we calibrate the status-quo economy that has no immigrants, to simulate the pre-shock equilibrium level of sustainable debt. To capture the impact of immigration on sustainable debt, we compare two economies, one which starts with residents only and the other one with immigrants with specific characteristics. To do so, we identify the factors through which immigration modifies the baseline economy and consider impulse responses from each of these factors, holding the remaining ones to what Canada immigration facts suggest. We run a set of experiments in which, a high skill immigrant is relatively less efficient than a high skill resident; we assume equal efficiency between low skill immigrants and low skill residents.[2]

Overall, our results suggest that immigration improves fiscal solvency. Not surprisingly, the impact of labor efficiency is the highest: the more efficient immigrants are, the more the host country can produce, the more revenues the government can make. Less obvious, our results suggest that there is an optimal level of efficiency

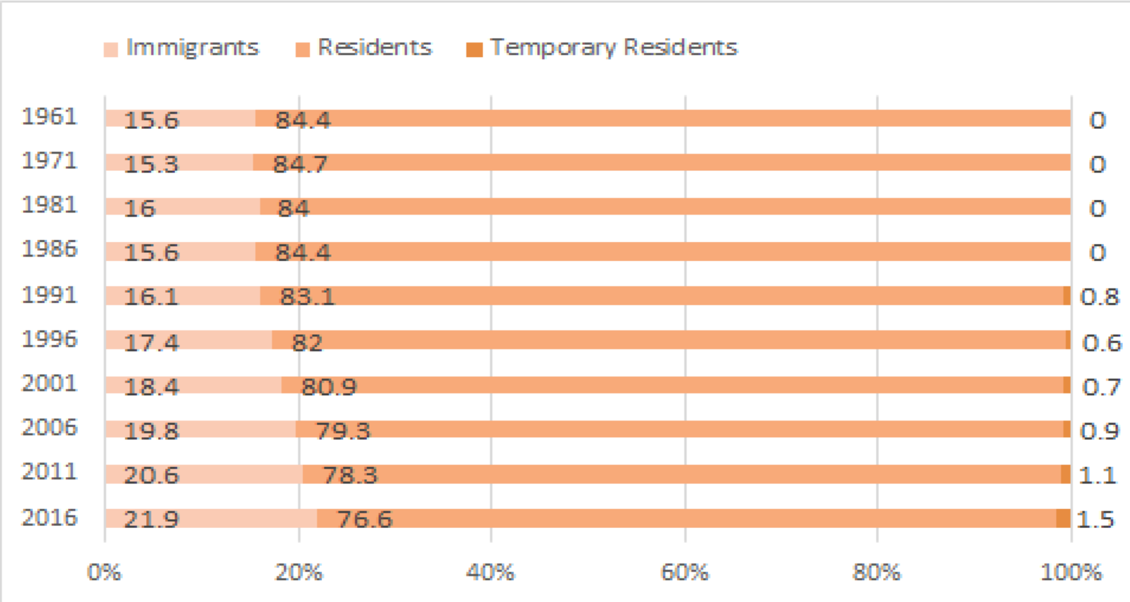
[2] In Canada, credentials obtained abroad are generally underscored. High skill immigrants undergo some training before entering the labor market.

to expect from immigrants. This is because at some point, if the taxation schedule is not appropriate, the equality in wealth distribution is jeopardized by the amount of wealth immigrants can make. The age structure of the immigrating population also matters first because taxation on labor occurs during the working life. We also find that fiscal solvency is positively affected by the share of kids in the overall immigrating population. The first reason is that kids don't affect government primary balances while they are kids. Indeed, they consume all the transfers they receive from the government. In addition, they have high probabilities to survive until the age they become productive through their entry in the labor market.

The positive impact of immigration on fiscal solvency is mostly driven by a positive change in public bond price and a slight increase of period-by period primary balances. Because the steady state share of the population that saves is higher, demand for public bond increases, therefore, public bond price increases, improving government fiscal solvency. We should note that these are results for a one-shot immigration shock. However, elaborate immigration policies will plausibly ensure a continuous influx of young immigrants that will keep supporting the previous generation. Our results are therefore lower bounds of immigration effects.

The remaining of the paper is structured as follows. In section 2.2 , we present the structural model used as framework to study immigration and sustainable debt. Section 2.3 is dedicated to the calibration of the economy with no immigrants; section 2.4 is dedicated to the calibration of the modified economy with immigrants. In section 2.5, we perform the set of our quantitative exercises and section 2.6 concludes.

Figure 2.2: A look at the evolution in the share of newly immigrants into Canada



Sources. From Authors based on Data from Statistics Canada.

2.2 Immigration and Sustainable debt: A Framework

Our framework nests the dynamic equilibrium model of d'Erasmo et al (2016) and Conesa & Garriga (2008). The main components are households' heterogeneity, life cycle with survival risk, efficient units of labor, fertility and social security through retirement pensions. Households' heterogeneity is captured by skill differences which are modeled through differentials in efficiency units of labor.[3]

The economy starts at date t_0. We consider a competitive equilibrium with heterogeneous households, a representative firm and the government. The households consume and supply labor to the firm. The firm produces using the labor supplied by households. The government redistributes the wealth levied through taxes on households and firms. Later, we discuss the objectives of each of these agents more extensively.

[3]It is also possible to inherently distinguish between skilled and unskilled labor, so that they enter differently in the production function. In such setup, there would be a specific price for each type of labor. We abstract from this specificity for now.

The main objective of the paper is to assess how the entry of immigrants affects the sustainability of the debt in the host country. Therefore, we will begin by studying the status quo which is the economy with no immigrants. This setup will also serve to calibrate the model for our quantitative exercises. We will then show how the framework is modified, once immigrants are added to the picture.

2.2.1 Baseline economy

In this section, I will discuss in detail the economy with no immigrant. We introduce immigrants in section 2.2.3 further below.

Households

The baseline economy consists of overlapping generations of resident consumers with stochastic lifetimes that last up to $\mathscr{I}$ years. We introduce immigrants in Section ... further. We denote the conditional probability of survival from age i to $i+1$ by ϕ_i. The unconditional probability of living until age i is then given by $s_i = \prod_{j=1}^{i-1} \phi_j$. Denoting γ_r, the population growth rate, the measure of households of age i at time t $\mu_{i,t}$, is computed as:

$$\mu_{i,t} = \phi_{i-1}\mu_{i-1,t-1} \quad with \quad \mu_{1,t} = (1+\gamma_r)\mu_{1,t-1}. \tag{2.1}$$

(3.1) implies that $\mu_{i,t} = (1+\gamma_r)\mu_{i,t-1}$ for any i. As the economy begins at time t_0, the number of resident consumers is given by $N_{t_0}^R = \sum_i^J \mu_{i,t_0}$.

Workers enter the labor market at age i_w and retire at age i_r. They have 1 unit of time to split between work and leisure. They differ in the skill content of their labor hours, which varies by their work experience, as captured by their age, i, and by their skill level $j \in \{L, H\}$. A fraction Φ of residents are high skilled (H) and the remainder is low skilled (L). Let ϵ_{ij} denote the effective labor supply per unit of time in efficiency units. Note that skill types differ in the age profiles of their productivity.[4]

[4]This fraction is the same at each period because survival probability is only age-dependent.

Households consume at every age. They have access to the credit market only after they enter the labor market, meaning at ages above i_w. Public bond d is the only asset.

Consumption and wealth accumulation are funded by resources drawn from different sources: earnings from labor during active life, return on wealth, and government transfers made of lump sum transfers e_t to all agents, and pensions p_t paid to retirees. Consumers also pay taxes to the government, specifically a tax on consumption τ_c and a tax on labor income τ_L.

In period t, the utility of a consumer with skill j, born in period s ($t - I < s <= t$) whose age is i=t-s, is given by $u(c_{i,j,t}, l_{i,j,t})$. Agents choose consumption stream c and labor supply l (rented to firms), as well as wealth transfer d to maximize lifetime utility subject to budget constraints. In our setup with life cycle and periods of inactivity and activity, households face three budget constraints throughout their lifetime. Before labor market entry, the consumer is inactive and does not have access to the credit market.

$$U = \sum_{i=1}^{i=\mathscr{I}} s_i \beta^i u(c_{i,j,t}, 1 - l_{i,j,t}) \tag{2.2}$$

The lifelong utility is the discounted sum of utility at each age i, the discount rate being β. Age i utility is also weighted by the survival rate s_i, which is the probability that the consumer reaches age i. Before working, meaning for i satisfying $i < i_w$:

$$(1 + \tau_c)c_{i,j,t} = e_t \tag{2.3}$$

The consumer pays a consumption tax τ_c. Before they are active workers, consumers are excluded from the financial market, so that the only revenue is the lump sum transfer from the government e. During the working life, meaning for i satisfying $i_w \leq i < i_r$:

$$(1 + \tau_c)c_{i,j,t} + (1 + \gamma)q_t d_{i,j,t+1} = (1 - \tau_L)w_t \epsilon_{i,j} l_{i,j,t} + d_{i,j,t} + e_t \tag{2.4}$$

Active workers accumulate wealth through public debt purchases d; the factor $(1 +$

γ) results from imposing balanced growth, with γ the growth rate of production.[5] The consumer of age i and skill j, that allocates $l_{i,j,t}$ units of time to work, earns $(1 - \tau_L)w_t\epsilon_{i,j}l_{i,j,t}$ after labor $\tau_L w_t\epsilon_{i,j}l_{i,j,t}$ taxes are levied.

After retirement, meaning for i such that $i_r \leq i \leq \mathscr{I}$

$$(1 + \tau_c)c_{i,j,t} + (1 + \gamma)q_t d_{i,j,t+1} = p_t + d_{i,j,t} + e_t \quad (2.5)$$

Unlike active workers, retirees do not earn labor revenue; instead, they receive pensions p from the government.

Firms

The representative firm rents labor from households and produces y_t.[6]

$$y_t = f(l_t) \quad (2.6)$$

Government

The government intervenes in the economy through outlays, taxes and public indebtedness. More specifically, revenues come from consumption taxes (τ_c), labor income taxes (τ_l) and debt issuance (d_t).[7] These revenues are allocated to public consumption g, lumpsum transfers to all consumers e and pension paid to retirees p, all taken exogeneously. In other words, the government primary balance pb_t, which is equal to revenues net of expenses, is funded by the change in debt net of debt service. We assume the government is committed to repay its debt, and thus it must satisfy the following sequence of budget constraints for t = t_0, ...,∞. We denote L_t aggregate

[5] Production growth rate is a function of the population growth rate.

[6] We abstract from capital. This is not essential for the purpose of the study, because we already have public debt as asset. Moreover, it is overall accepted that there is a rate to which capital can be substituted with labor.

[7] We consider tax on consumption and labor income as in D'Erasmo et al. (2016) but exclude tax on savings because we chose to keep savings free of distortions. Consumption on overall income would have achieved the same.

labor, C_t aggregate consumption, and P_t, aggregate pension paid.

$$pb_t = \tau_C C_t + \tau_L w_t L_t - (g_t + E_t + P_t) \tag{2.7}$$

$$pb_t = d_t - (1+\gamma) q_t d_{t+1} \tag{2.8}$$

With:

$$L_t = \sum_{s=t-\mathscr{I}}^{s=t} (\Phi \mu_{t-s,t} \epsilon_{t-s,H,t} l_{t-s,H,t} + (1-\Phi) \mu_{t-s,t} \epsilon_{t-s,L,t} l_{t-s,L,t})$$

$$C_t = \sum_{s=t-\mathscr{I}}^{s=t} (\Phi \mu_{t-s,t} c_{t-s,H,t} + (1-\Phi) \mu_{t-s,t} c_{t-s,L,t})$$

$$E_t = (\sum_{s=t-\mathscr{I}}^{s=t} \mu_{t-s,t}) e(t)$$

$$P_t = (\sum_{s=t-i_r}^{s=t} \mu_{t-s,t}) p(t)$$

Following d'Erasmo et al (2016), **public debt is sustainable** if the Intertemporal Government Budget Constraint (**IGBC**) holds. The IGBC condition is equivalent to the government satisfying a No-ponzi game condition: the discounted value of the stream of primary fiscal balances equals the initial public debt d_0. When the model is worked in shares of GDP, y_t, which will be the case for model calibration, the IGBC in shares of GDP writes:

$$\frac{d_{t_0}}{y_{t_0-1}} = \frac{y_{t_0+1}}{y_{t_0}} (\frac{pb_{t_0}}{y_{t_0}} + \sum_{t=1}^{\infty} ([\prod_{i=0}^{t-1} v_i] \frac{pb_t}{y_t})) \quad with \quad v_i = (1+\gamma) \frac{y_{i+1}}{y_i} \tag{2.9}$$

2.2.2 Equilibrium

In our quantitative analysis, we study the recursive competitive equilibrium of the economy defined formally below. **Proposition** *Given preferences, initial population structure* $(N^R_{t_0}, \{\mu_{i,t_0}\}_{i=1}^{\mathscr{I}}, \Phi, \gamma_r)$ *and taxation schedule* (τ_C, τ_L)*, an equilibrium is a collection of allocations for high skill resident* $\{c_{i,H,t},\ l_{i,H,t},\ d_{i,H,t+1}\}_{i=1...\mathscr{I}, t=t_0...+\infty}$*, low skill* *resident* $\{c_{i,L,t},\ l_{i,L,t},\ d_{i,L,t+1}\}_{i=1..\mathscr{I},t=t_0...+\infty}$*, a demand schedule from the firm* $\{l_t\}_{t=t_0}^{+\infty}$*, a government policy* $\{E_t,\ P_t,\ g_t\ d_{t+1}\}_{t=t_0}^{+\infty}$*, and a price system* Q=$(q_t)_{t=t_0}^{+\infty}$ *such that the following is satisfied:*

i *Optimality: given the price system Q, consumers' utility and firms' profit are maximized.*[8]

ii *Feasibility: the market for good, the market for labor, and the market for public debt clear for all t:*

$$\text{(Good) } Y_t = C_t + g_t \tag{2.10}$$

$$\text{(Labor) } l_t = L_t \tag{2.11}$$

$$\text{(Public Debt) } d_t = \sum_{s=t-\mathscr{I}}^{s=t} (\Phi\mu_{t-s,t}d_{t-s,H,t} + (1-\Phi)\mu_{t-s,t}d_{t-s,L,t}). \tag{2.12}$$

iii *The government's policy satisfies its budget constraint:*
$pb_t = \tau_C C_t + \tau_L w_t L_t - (g_t + E_t + P_t) = d_t - (1+\gamma_r)q_t d_{t+1}$.

The next section discusses how the entry of immigrants modified our baseline economy.

2.2.3 Adding immigrants to the economy

The main objective is to evaluate how the entry of immigrants affects sustainable debt. It is therefore critical to study the population dynamics for the quantitative part. We add upper-scripts to differentiate immigrants (i) to residents (r).

[8]Optimality conditions are provided in appendix.

Demographic structure and dynamics

$(N^R_{t_0}, \{\mu_{i,t_0}\}_{i=1}^{\mathscr{I}}, \Phi, \gamma_r)$ characterizes the baseline population structure. Let's consider the modified economy where there is a mass $N^I_{t_0}$ of immigrants with a proportion λ that are skilled. $\gamma_{r,i}$ is the adjusted population growth rate so that $(N^R_{t_0} + N^I_{t_0}, \{\mu^R_{i,t_0} + \mu^I_{i,t_0}\}_{i=1}^{\mathscr{I}}, \Phi\lambda, \gamma_{r,i})$ characterizes the new demographic structure. The initial structure evolves over time due to death probability and different fertility rates between immigrants and residents. We classify all newly born consumers are residents. As Woldmicael and Roderic (2010) show, fertility is on average higher for immigrants entering Canada, as compared to canadian-born, so that $\gamma_{r,i} > \gamma_r$ until all reproductive immigrants present at t_0 disappear. Then $\gamma_{r,i} = \gamma_r$.
The share of immigrant households in total population N_t, is given by $\eta_t = \frac{N^I_t}{N_t}$. The share of residents in total population N_t, is given by $1 - \eta_t = \frac{N^R_t}{N_t}$. Since survival rate is the same among resident and immigrant consumers, having all newly born as residents implies that the growth rate of the resident population is the same as the growth rate of the total population. Therefore the share of the residing population remains constant over time: $1 - \eta_t = 1 - \eta_{t-1}$ so $\quad \eta_t = \eta$.

Next, we characterize the immigrant consumer taking the resident consumer as reference.

Differences between residents and immigrants

In addition to fertility differentials, the main difference between residents and immigrants relates to labor efficiency. In the quantitative part, most of our experiments are performed with the assumption that high skilled residents are more efficient than high skilled immigrants, while low skilled immigrants as efficient as low skilled residents.[9] We also assess the sensitivity of sustainable debt to the parameter capturing the relative labor efficiency of low skill immigrants.

[9] High skill workers will generally go through some training before integrating the labor market.

So:

$$\epsilon^r_{i,H} > \epsilon^i_{i,H} \quad we \quad set \quad \epsilon^i_{i,H} = \kappa \epsilon^r_{i,H} \quad with \quad \kappa < 1$$
$$\epsilon^i_{i,L} = \psi \epsilon^r_{i,L} \quad we \quad set \quad \psi = 1 \quad in \quad most \quad experiments.$$

In the next section, we discuss the calibration of the baseline economy.

2.3 Calibration of Baseline Economy to Canada

2.3.1 Demographics

We set the length of life to three periods, so the household spends one period in each phase of life: childhood, employment and retirement. Thus, $i_w = 2$ and $i_r = 3$. One period would be 30 years so that households are inactive from 0 to 30 years old , they work from 30 and retire at 60 years of age.[10] Survival probabilities ϕ_i are taken from Bell and Miller (2005) and aggregated to match the age profile in our setup. This implies survival probability of 0.996 from age 1 to 2, of 0.953 from age 2 to 3 and of 0 at age 3.

From Woldemicael and Beaujot (2010), we set the average number of residents' children to 1.59.[11] Assuming that only age 2 agents reproduce, $\gamma_r = 0.58$ as shown

[10]The assumption of three periods implies that the working life and retirement are of the same length. Thus, the dependency ratio implied by our model would be larger than in reality. Having longer working life would improve public finance and thus debt sustainable irrespective of the presence of immigrants. It could be interesting to assess how sensitive debt sustainability is to the length of working life.

[11]Woldemicael and Beaujot (2010) estimate to 1.76 the number of children from 35-44 years old foreign born women in 2002, and to 1.59 the same for canadian-born women of the same ages, for the same year.

below:[12]

$$\mu_{1,t} = 1.59\mu_{2,t}$$
$$\mu_{2,t} = \phi_1 \mu_{1,t-1}$$
$$\mu_{1,t} = 1.59\phi_1 \mu_{1,t-1}$$
$$\gamma_r = \frac{\mu_{1,t} - \mu_{1,t-1}}{\mu_{1,t-1}} = 1.59\phi_1 - 1 = 0.58$$

Using Population data from Canada in 2018, we set the initial population age profile to $(\mu_{1,t_0}, \mu_{2,t_0}, \mu_{3,t_0}) = (0.16, 0.66, 0.17)$, which implies that $N^R_{t_0} = 1$.[13] We set the share of high skill workers to 0.4.[14]

2.3.2 Endowments

Efficient units of labor is households' endowments. Our efficiency units of labor are based on Hansen (1993)'s estimates who provide efficiency units for a finer age profile. Thus, we average to get efficient units of labor in our case, assuming that the values provided are those of a highly skilled worker. Since workers are active for only one period that covers ages from 30 to 60, we get $\epsilon^R_{2,H} = 1.97$. To get efficient units of labor for low skilled workers, we use hourly wage by union coverage status: workers with no union coverage earn 83% of the hourly wage of unionized workers. Thus, $\epsilon^R_{2,L} = 1.97 \times 0.83 = 1.64$. For Consumers also receive lump-sum transfers e_t at all ages; pensions p_t are paid to retirees. We impose that both type of transfers are fixed over time. We provide their values below, as we discuss the calibration of the public sector.

[12] One period is equivalent to 30 years in this model; this explains the high implied population growth.

[13] The share by age groups are taken from Annual Demographic Estimates: Canada, Provinces and Territories, 2018, Statistics Canada, Demography Division.

[14] Figure 1, OECD (2004), for 2001.

2.3.3 Government

Taxes are taken directly from OECD releases: we set consumption tax to $\tau_C = 0.05$ and labor income tax to $\tau_L = 0.19$.[15]

To compute variables related the government, we take some directly from data provided by Statistics Canada (the fourth quarter of 2019). For government final consumption g, we impute $\frac{g}{Y} = 0.20$.[16] Primary balance as share of GDP is set to $\frac{Pb}{Y} = 0.01$.[17] To get the implied value of total transfers to households $\frac{E+P}{Y}$, we need $\frac{C}{Y}$ which is obtained from the equilibrium on the market for final good $\frac{C}{Y} = 1 - \frac{g}{Y} = 0.8$, so that $\frac{E+P}{Y} = 0.04$. From OECD stats, $\frac{P}{Y} = 0.048$ which exceeds the value that our model would imply, thus, we set $\frac{P}{Y}$ to 0.03 and $\frac{E}{Y}$ to 0.01.[18] Turning to the value of transfers per consumer consumer, since $N_{t_0}^R = 1$, e=E; e is fixed over time, so $\frac{E}{Y}$ changes over time. Pensions are paid to retirees only, so $\frac{p}{Y} = \frac{\frac{P}{Y}}{\mu_{t_0}} = 0.18$.

2.3.4 Functional Forms

Households preferences are assumed to take the following form: $u(c, l) = logc + log(1 - l)$ which implies a relative risk aversion of 1.

So, for j={L,H}, $U_j = \sum_{i=1}^{i=3} s_i \beta^i (log(c_{i,j,t}) + log(1 - l_{i,j,t}))$.

As of technology, we work with the following production function $f(L_t) = L_t$ so that wage is equal to 1 at each period.

The table below resumes how we calibrated the status-quo economy (without immigrants).

---

[15] OECD Stats, Table I.6. All-in average personal income tax rates at average wage by family type.

[16] Authors' computations based on Statistics Canada. Table 36-10-0222-01 Gross domestic product, expenditure-based, provincial and territorial, annual (x 1,000,000) is used to compute consumption as shares of GDP.

[17] Authors' computations based on Statistics Canada. Table 36-10-0477-01 Revenue, expenditure and budgetary balance - General governments (x 1,000,000).

[18] OECD (2021), Pension spending (indicator). doi: 10.1787/a041f4ef-en (Accessed on 30 June 2021).

Table 2.1: Calibration of baseline economy

Parameters	Value	Source
γ_r	0.58	Authors computations based on Woldmicael et al. (2010)
i_w	2	From authors
i_r	3	From authors
$(\mu^R_{1,t_0}, \mu^R_{2,t_0}, \mu^R_{3,t_0})$	(0.16,0.66,0.17)	Canada population Data (2018)
(ϕ_1, ϕ_2, ϕ_3)	(0.99, 0.95, 0)	Averages based on estimates from Bell and Miller (2005)
Φ	0.4	Figure 1, OECD (2004), for 2001
$\epsilon^R_{2,H}$	1.97	Hansen (1993)
$\epsilon^R_{2,L}$	1.64	Hourly wage ratio (no union coverage to union coverage), 2019
(τ_C, τ_L)	(0.05,0.18)	OECD Stats
β	0.998	D'Erasmo et al. (2016)

Notes. We use the hourly wage ratio between workers uncovered by a union and workers covered by a union to capture the relative efficiency between high skill and low skill workers.

Next, we show how key parameters of the baseline economy are affected by the entry of immigrants.

2.4 Economy with immigrants

Some components of the population structure are modified. At t_0, the population structure is now $(N^R_{t_0} + N^I_{t_0}, \{\mu^R_{i,t_0} + \mu^I_{i,t_0}\}^{\mathscr{I}}_{i=1}, \Phi, \lambda, \gamma_{r,i})$. Based on data from Statistics Canada and as shown on Figure 2.2, newly arrived immigrants represent approximately 22% of the whole population of Canada. Thus

$$N^I_{t_0} = 0.22(N^I_{t_0} + N^R_{t_0})$$
$$N^I_{t_0} = 0.28$$

. The age profile of immigrants in 2011 implies that $(\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0}) = (0.33, 0.64, 0.03)$.[19] As of the share of highly skilled immigrants, following results of King (2009), we set

[19] Sources. Statistics Canada, Catalogue no. 99-010-X2011001 ISBN: 978-1-100-22197-7.

λ to 0.41 .[20] Based on Woldemicael and Beaujot (2010), the number of children from 35-44 years old foreign born women in 2002 is 1.76. Thus, the growth rate at t_0 is given by

$$\gamma_{r,i} = \frac{1.59\mu^R_{2,t_0} + 1.76\mu^I_{2,t_0} - \mu^R_{1,t_0} - \mu^I_{1,t_0}}{\mu^R_{1,t_0} + \mu^I_{1,t_0}}$$

$$\gamma_{r,i} = 3.43$$

From $t_0 + 1$ on, there is no age 2 immigrants because all newly born are residents; thus $\gamma_{r,i} = 0.58$.

To get the relative labor efficient units between immigrants and residents, we use weekly wages of university educated new immigrants and compare it to their resident counterparts. In 2006, university graduated immigrants earned on average 88% of the weekly wage of university graduated canadian born.[21] So $\epsilon^I_{2,H}$=1.97$\times$0.88 $= 1.75$.

[20] Martin Prosperity Institute REF. 2009-WPONT-012.

[21] Source(s): Canadian censuses of 1981, 1991, 2001 and 2006 20% files; U.S. censuses of 1980, 1990, and 2000 IPUMS 5% files and 2005 American Community Survey IPUMS 1% file.

Table 2.2: Calibration of modified economy

Parameters	Value	Source
$\gamma_{r,i}$	(3.43; 0.58)	Authors computations based on Woldmicael et al. (2010)
i_w	2	From authors
i_r	3	From authors
$(\mu^R_{1,t_0}, \mu^R_{2,t_0}, \mu^R_{3,t_0})$	(1, 0.16, 0.66, 0.17)	Canada population Data
$(\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0})$	(0.33, 0.64, 0.03)	Canada population Data (2011)
N^I_{1,t_0}	0.28	Computations based on Figure 2.2, 2016
(ϕ_1, ϕ_2, ϕ_3)	(0.99, 0.95, 0)	Averages based on estimates from Bell and Miller (2005)
λ	0.41	King (2009), value for 2006
$\epsilon^I_{2,H}$	1.75	Hansen (1993) and Statistics Canada, 2006
$\epsilon^I_{2,L}$	1.64	Weekly wage university graduates, 2006
(τ_C, τ_L)	(0.05,0.18)	OECD Stats
β	0.998	D'Erasmo et al. (2016)

Notes. We use the hourly wage ratio between workers uncovered by a union and workers covered by a union to capture the relative efficiency between highly skilled and low skill workers.

We perform a set of exercises to assess the impact of receiving immigrants on government fiscal solvency. The parameters that capture how immigration affects the economy are $(\lambda, \epsilon^I_{2,H}, \epsilon^I_{2,L}, 1-\eta, \gamma_{r,i}, (\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0}))$. In all the experiments we conduct below, the entry of immigrants is a temporary, meaning that in the modified economy, a mass $N^I_{t_0}$ of immigrants enter at date t_0. Also, all newly born are residents. In the next section, we provide our quantitative results on the effect of immigration on sustainable debt.

2.5 Quantitative assessment of the effect of immigration on sustainable debt

Firstly, we assess how the share of immigrants in the whole population affects debt sustainability.

2.5.1 Sensitivity to the total share of immigrants $1-\eta$

Here, we are assessing how fiscal solvency is affected by the total share of immigrants in the population. Therefore, we compute sustainable debt as a function of that share. We use Canada immigration data to set a value to the remaining parameters $(\lambda, \epsilon^I_{2,H}, \epsilon^I_{2,L}, \gamma_{r,i}, (\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0})) = (0.41, 1.75, 1.64, (3.34, 0.58), (0.33, 0.64, 0.03))$. Using these values, we get the sustainable debt, which is the present discounted value of inter-temporal primary balances. With status-quo sustainable debt as baseline value (without immigrants), we compute the percentage change: the variable along the vertical axis. So the curve on figure 2.3 maps values of $1-\eta$ into the percentage change of sustainable debt: the equilibrium present discounted value of the primary fiscal balance for the modified economy (with immigration) relative to the simulated value for the baseline economy (without immigration).

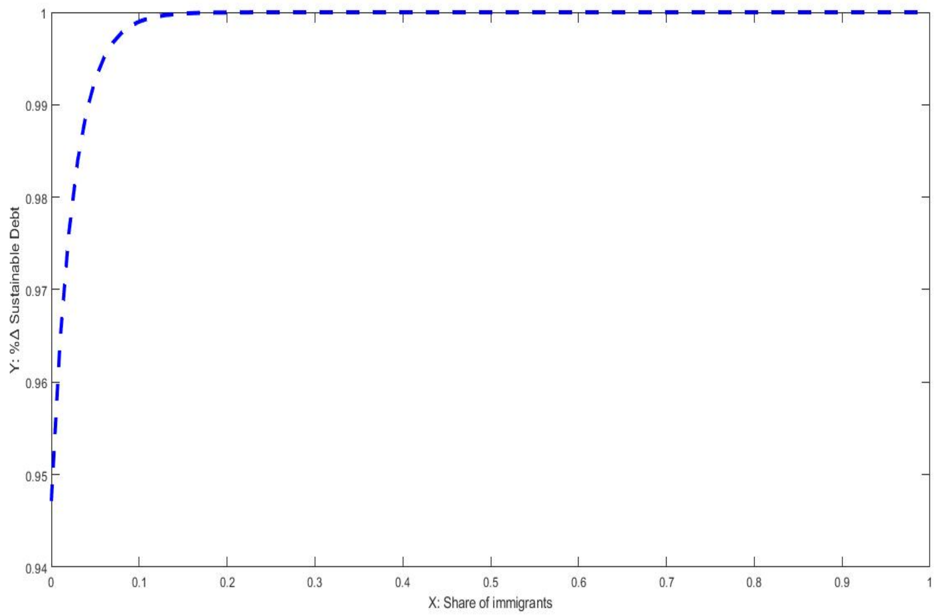


Figure 2.3: Change in sustainable debt as a function of immigration intensity

We see that the entry of immigrants increases sustainable debt, since the values on the vertical axis are all positive. There is a sharp increase in sustainable debt for low values of the mass of entrants, then the change stagnates to 1, which means that sustainable debt has doubled. This happens once we hit a share of immigrants of approximately 15%. This result was predictable because what matters the most is the age profile of immigrants, which is fixed to the age profile of canadian immigrants. Indeed, each immigrant pays consumption tax at all ages, pays income tax only during working life and retirement, receives lumpsum transfers at all ages and pensions during retirement. When immigrants enter, it creates a shock to the status quo economy and primary balance would be positively affected if the share of active immigrants is high enough. It is important to note immigration data from Canada features a high fraction of immigrants in their working life. It is also interesting to note that each year, the share of immigrants in the whole population is approximetely 20%.

2.5.2 Sensitivity to age profile of immigrants: fractions taken by pairs $(\mu^I_{1,t_0}, \mu^I_{3,t_0})$, $(\mu^I_{2,t_0}, \mu^I_{3,t_0})$

To better capture the influence of the age profile, it takes to consider fractions of the population at least by pairs. Indeed these fractions are dependent from each other since $(\mu^I_{1,t_0}+\mu^I_{2,t_0}+\mu^I_{3,t_0} = 1)$. So, we compute sustainable debt as a function of two ages. As before, we use Canada immigration data provided in table 2.2 to set a value to the remaining parameters $(\lambda, \epsilon^I_{2,H}, \epsilon^I_{2,L}, 1-\eta, \gamma_{r,i}) = (0.41, 1.75, 1.64, 0.28, (3.43, 0.58))$. The curve on the left of figure 2.4 maps $(\mu^I_{1,t_0}, \mu^I_{3,t_0})$ to the percentage change of sustainable debt, and the second graph maps changes in $(\mu^I_{2,t_0}, \mu^I_{3,t_0})$ to the percentage change in sustainable debt. Thus, the younger the immigrating population, the more the the public finances of hosting country are improved.

2.5.3 Sensitivity to immigrants' skill $(\epsilon^I_{2,H})$

The focus here is on the impact on fiscal solvency of the relative efficiency in labor of immigrants. We set the remaining parameters from data $(\lambda, 1-\eta, \epsilon^I_{2,L}, \gamma_{r,i}, (\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0})) =$

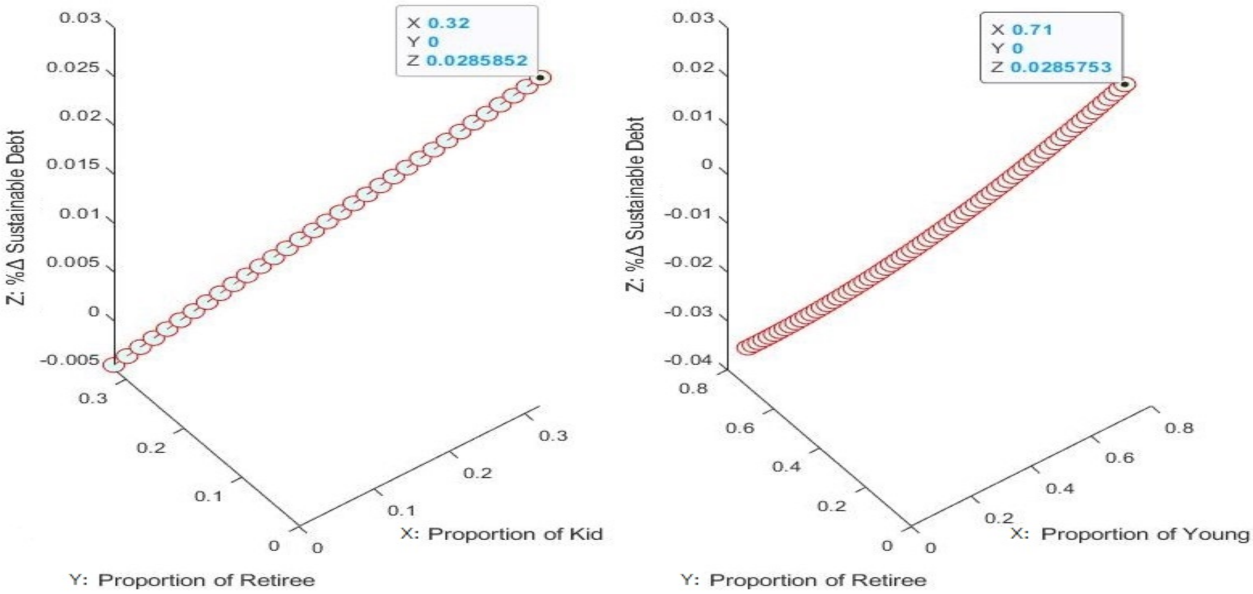


Figure 2.4: Change in sustainable debt as a function of immigrants' age fractions taken by pairs

$(0.41, 0.28, 1.64, (3.43; 0.58); (0.33, 0.64, 0.03))$. Figure 2.5, maps $\epsilon_{2,H}^{I}$ to the change in sustainable debt once immigration has happened. Thus, the more immigrants are efficient in labor, the better debt sustainability would be after they enter the country. However, above a certain threshold, the gain in debt sustainability decreases with the labor efficiency of immigrants. In fact, consumption of active workers is a function of their labor efficiency. However, tax on labor income is flat which might create distribution issues. Indeed, consumption of kids are fully funded by government transfers and consumption retirees are partially funded by pensions paid by the government. At some point, public revenues does not increase as much as immigrants wealth does, leading to a relatively lower gain in debt sustainability.

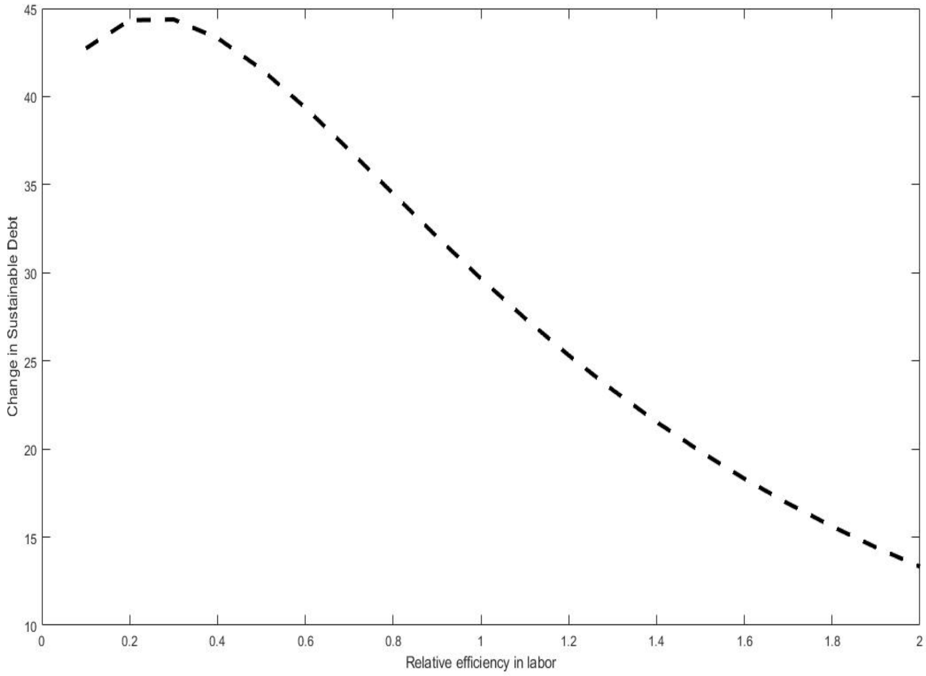


Figure 2.5: Change in sustainable debt as a function of immigrants' labor efficiency and type of skill intensity

2.5.4 The reason behind: Immigration shock, primary balance and price of public bond

From results just discussed, immigration appears to be beneficial for the host country fiscal solvency. We saw that sustainable debt is the discounted present value of all primary balances; it is equivalent to the no-ponzi game condition for the government. Therefore if one can measure the effect of immigration on asset price (inverse of gross return) and on each period primary balance, it will be straightforward to grab the change in sustainable debt. The figure below addresses this point. Immigration parameters are set like this $(\lambda, \epsilon^I_{2,H}, \epsilon^I_{2,L}, 1-\eta, \gamma_{r,i}, (\mu^I_{1,t_0}, \mu^I_{2,t_0}, \mu^I_{3,t_0})) = (0.41, 1.75, 1.64, 0.28, (3.34, 0.58), (0.33, 0.64, 0.03))$. The graph on the left represents two curves, primary balance as shares of GDP from period $t = t_0 = 0$ to period t=30, for the baseline economy (without immigrants in blue) and for the modified economy (after immigration, in red). The graph on the right does the same for asset (public bond) price.

An immigration shock similar to Canada yearly immigration will initially deteriorate government primary balance. It is not clear on the graph, but at some point, primary balance becomes positive and larger than what it would have been in the absence of immigrants. For asset price, it takes some years before the equilibrium price of public bond differs from what it would have been without immigrants; we can see that the gap is pretty large. Both the delay and the direction of the change in the asset price are consistent with what the model predicts. In fact, since the pool of immigrants is mostly made of active workers, the immigration shock increases significantly the share of households that are savers. With more savings, the demand for public bonds increases, leading to a higher price. The delay is consistent with the fact that it takes some time for assets to accumulate, and for the debt stock to significantly change.

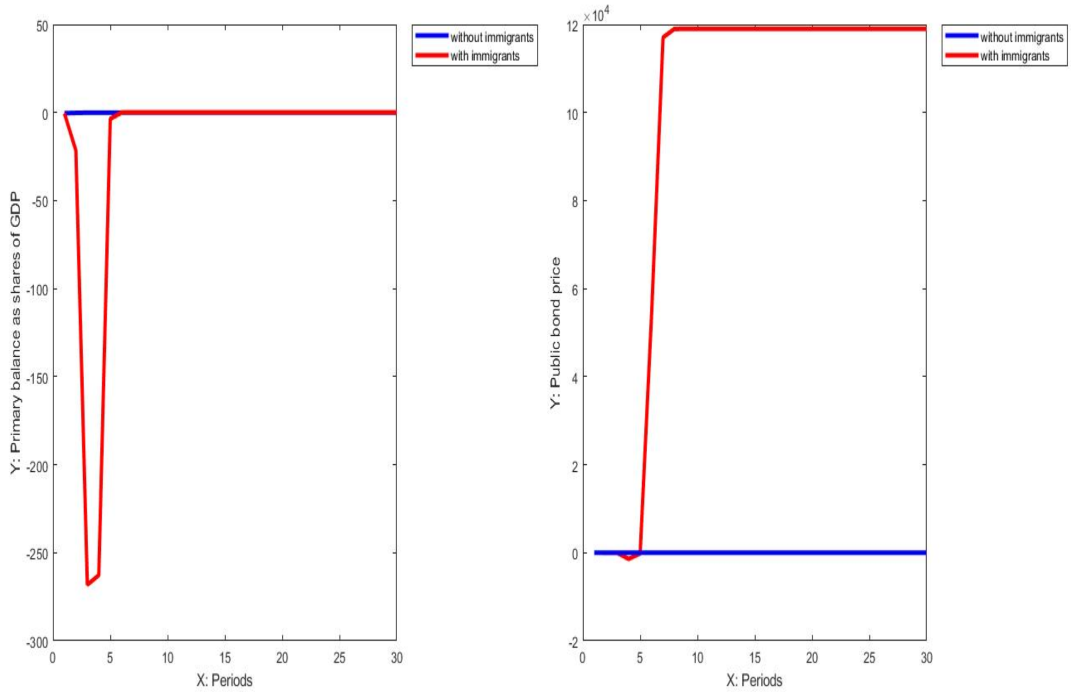


Figure 2.6: Primary balance and public bond price over time

2.6 Conclusion

Countries that receive immigrants have their population structure significantly modified. Indeed, immigration is a particular kind of demographic shock, since immigration "newborns" may have a past and characteristics that differ from natives: their skills and their propensity to reproduce are examples of those. Aging societies as Canada generally undergo massive immigration.

This paper addressed one aspect, and not the least of the various impacts that immigration has on the hosting country: fiscal solvency. When immigrants enter a country, they work, they consume, they save, all things that are beneficial for the receiving country. But they will also get ill, age, loose jobs, which will require government support. The fact that aging societies continue to receive immigrants suggests that there is more to gain that to loose. Our paper proposed a rationale to explain the positive impact of immigration. Using a DSGE model that features life cycle, death risk, skills and labor efficiency, we simulated the equilibrium effect of immigration on sustainable debt.

Sustainable debt is the level of debt that all subsequent primary balances will cover with strict equality. So the more a country is able to make fiscal surpluses, the more room it has to borrow while remaining solvable. There are different dimensions by which immigration changes the population characteristics, and whatever the aspect we consider, our results show that immigration improves fiscal solvency. The entry of immigrants, mostly made of active workers increases the share of the population that saves. Therefore, the demand for public bonds increases driving price of public bond up. Age and skills of immigrants are characteristics that boost fiscal solvency the most. The younger the immigrating population, the better the public finances. The more efficient the immigrating population is, the better the fiscal solvency of the host country. However, with a flat tax rate on labor income, at some point, the positive impact of immigrants' labor efficiency on the fiscal solvency is reduced. The reason is that government revenues do not increase as much as immigrants' revenues. This result emphasizes that immigration may affect wealth distribution.

Bibliography

Altonji, Joseph G. and Charles Pierret, "Employer learning and statistical discrimination," *Quarterly Journal of Economics*, 1997, *116* (2), 313–350.

Bauer, Thomas K and John P Haisken-DeNew, "Employer learning and the returns to schooling," *Labour Economics*, 2001, *8* (2), 161–180.

Bell, Felicitie C. and Michael L. Miller, "Life tables for the United States social security area 1900–2100," *Social Security Administration, Office of the Chief Actuary, SSA*, 2005.

Conesa, Juan C. and Carlos Garriga, "OPTIMAL FISCAL POLICY IN THE DESIGN OF SOCIAL SECURITY REFORMS*," *International Economic Review*, 2008, *49* (1), 291–318.

D'Erasmo, P., E.G. Mendoza, and J. Zhang, "Chapter 32 - What is a Sustainable Public Debt?," in John B. Taylor and Harald Uhlig, eds., *John B. Taylor and Harald Uhlig, eds.*, Vol. 2 of *Handbook of Macroeconomics*, Elsevier, 2016, pp. 2493–2597.

Farber, Henry S. and Robert Gibbons, "Learning and Wage Dynamics*," *The Quarterly Journal of Economics*, 11 1996, *111* (4), 1007–1047.

Galindo-Rueda, Fernando, "Employer learning and schooling-related statistical discrimination in Britain," *Journal of Human Resources*, 1993, *4*, 759–790.

Gibbons, Robert and Lawrence F. Katz, "Layoffs and Lemons," *Journal of Labor Economics*, 1991, *9* (4), 351–380.

Hansen, Gary, "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked," *Journal of Applied Econometrics*, 1993, *8* (1), 71–80.

Kahn, Lisa B., "Asymmetric Information between Employers," *American Economic Journal: Applied Economics*, October 2013, *5* (4), 165–205.

_ **and Fabian Lange**, "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures," *The Review of Economic Studies*, 07 2014, *81* (4), 1575–1613.

Kaymak, Barış, "Postschooling Training Investment and Employer Learning," *Journal of Human Capital*, 2014, *8* (3), 318–349.

King, Karen M., "The Geography of Immigration in Canada: Settlement, Education, Labour Activity and Occupation Profiles," *Martin Prosperity Institute*, 2009.

Krieger, Tim, "Fertility Rates and Skill Distribution in Razin and Sadka's Migration-Pension Model: A Note," *Journal of Population Economics*, 02 2004, *17*, 177–182.

Lange, Fabian, "The Speed of Employer Learning," *Journal of Labor Economics*, 2007, *25* (1), 1–35.

Lee, Ronald and Timothy Miller, "Immigration, Social Security, and Broader Fiscal Impacts," *American Economic Review*, May 2000, *90* (2), 350–354.

Pinkston, Joshua C., "A Test of Screening Discrimination with Employer Learning," *ILR Review*, 2006, *59* (2), 267–284.

Rowthorn, Robert, "The fiscal impact of immigration on the advanced economies," *Oxford Review of Economic Policy*, 10 2008, *24* (3), 560–580.

Schönberg, Uta, "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 2007, *25* (4), 651–691.

Storesletten, Kjetil, "Sustaining Fiscal Policy through Immigration," *Journal of Political Economy*, 2000, *108* (2), 300–323.

Woldemicael, Gebremariam and Roderic Beaujot, “Fertility Behavior of Immigrants in Canada: Converging Trends,” *PSC Discussion Papers Series*, 01 2010, *13*.

.1 Appendix to Chapter 1

.1.1 Optimal incumbent wage at t=2, w_2^*

The worker chooses to stay with the incumbent if $w_2 - \theta_2 > v_2$. The probability of staying is denoted $P_{accepts}(t)$ with θ_2 following the uniform distribution over the interval $[-\bar{\theta}; \bar{\theta}]$.

w_2^* maximizes $V_2(z) = (z - w_2)P_{accepts}(2)$, with

$$P_{accepts}(2) = \begin{cases} 0 & \text{if} \quad w_2 - v_2 < -\bar{\theta} \\ \frac{w_2 - v_2 + \bar{\theta}}{2\bar{\theta}} & \text{if} \quad w_2 - v_2 \in [-\bar{\theta}; \bar{\theta}] \\ 1 & \text{if} \quad w_2 - v_2 > \bar{\theta} \end{cases}$$

The incumbent firm maximizes the expected profit , taking outside firms behavior as given. Denoting the equilibrium outside wage offer as w_2^*, we see that there would be two cases: one interior solution when $w_2 - v_2 \in [-\bar{\theta}; \bar{\theta}]$ and one corner solution which is $w_2^* = v_2 + \bar{\theta}$, when $w_2 - v_2 > \bar{\theta}$. The corner solution yields no rent for the incumbent. In fact, the objective function is strictly quasi-concave when $w_2 - v_2 \in [-\bar{\theta}; \bar{\theta}]$

$$\begin{aligned} v_2(z) = max_{w_2}(z - w_2)\frac{w_2 - v_2 + \bar{\theta}}{2\bar{\theta}} \\ = max_{w_2}(\frac{zw_2 - zv_2 + z\bar{\theta}}{2\bar{\theta}} - \frac{w_2^2 - w_2v_2 + w_2\bar{\theta}}{2\bar{\theta}}) \\ = max_{w_2}\frac{-w_2^2 + w_2(z + v_2 - \bar{\theta}) + z(\bar{\theta} - v_2)}{2\bar{\theta}} \end{aligned}$$

First order conditions: $-2w_2^* + z + v_2 - \bar{\theta} = 0$, so $w_2^* = \frac{z + v_2 - \bar{\theta}}{2}$

.1.2 Optimal incumbent wage at t=1, w_1^*

w_1^* maximizes $V_1(\tilde{z}) = E[(z - w_1 + E(V_2))P_{accepts}(1)|\tilde{z}, s]$

$$P_{accepts}(1) = \begin{cases} 0 & \text{if} \quad w_1 - v_1 < -\bar{\theta} \\ \frac{w_1 - v_1 + \bar{\theta}}{2\bar{\theta}} & \text{if} \quad w_1 - v_1 \in [-\bar{\theta}; \bar{\theta}] \\ 1 & \text{if} \quad w_1 - v_1 > \bar{\theta} \end{cases}$$

The corner solution is the same as at t=2. The interior solution maximizes:

$$\begin{aligned} V_1(\tilde{z}) &= (\tilde{y}_1 + E(V_2|\tilde{z}, s) - w_1)(\frac{w_1 - v_1 + \bar{\theta}}{2\bar{\theta}}), \text{ with } \tilde{y}_1 = E(z|\tilde{z}, s) \\ = max_{w_1}(\frac{(\tilde{y}_1 + E(V_2|\tilde{z}, s))w_1 - (\tilde{y}_1 + E(V_2|\tilde{z}, s))v_1 + (\tilde{y}_1 + E(V_2|\tilde{z}, s))\bar{\theta}}{2\bar{\theta}} &- \frac{w_1^2 - w_1 v_1 + w_1\bar{\theta}}{2\bar{\theta}}) \\ = max_{w_1} &\frac{-w_1^2 + w_1(\tilde{y}_1 + E(V_2|\tilde{z}, s) + v_1 - \bar{\theta}) + (\tilde{y}_1 + E(V_2|\tilde{z}, s))(\bar{\theta} - v_1)}{2\bar{\theta}} \end{aligned}$$

First order conditions:

$$\begin{aligned} -2w_1^* + \tilde{y}_1 + E(V_2|\tilde{z}, s) + v_1 - \bar{\theta} = 0, \ so \\ w_1^* = \frac{\tilde{y}_1 + E(V_2|\tilde{z}, s) + v_1 - \bar{\theta}}{2} \end{aligned}$$

.1.3 Case 1: Increasing Asymmetry

In the case of increasing asymmetry, the information set of outside firms remains the same in all employment periods; they only observe schooling.

Optimal Outside Option at t=2, v_2^*

$$
\begin{aligned}
v_2^*(s) &= E(z|s, worker\ rejects\ w_2^*) \\
v_2^*(s) &= \frac{\int_{-\infty}^{\infty} z(1-P_{accepts}(2))\pi(z|s)}{\int_{-\infty}^{\infty}(1-P_{accepts}(2))\pi(z|s)} \\
v_2^*(s) &= \frac{\int_{-\infty}^{\infty} z(1-\frac{w_2-v_2+\bar{\theta}}{2\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} 1-\frac{w_2-v_2+\bar{\theta}}{2\bar{\theta}}\pi(z|s)} \\
v_2^*(s) &= \frac{\int_{-\infty}^{\infty} z(1-\frac{\frac{z+v_2-\bar{\theta}}{2}-v_2+\bar{\theta}}{2\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} 1-\frac{\frac{z+v_2-\bar{\theta}}{2}-v_2+\bar{\theta}}{2\bar{\theta}}\pi(z|s)} \\
v_2^*(s) &= \frac{\int_{-\infty}^{\infty} z(\frac{v_2-z+3\bar{\theta}}{4\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} \frac{v_2-z+3\bar{\theta}}{4\bar{\theta}}\pi(z|s)} \\
v_2^*(s) &= \frac{v_2^*(s)E(z|s)-E(z^2|s)+3\bar{\theta}E(z|s)}{v_2^*(s)-E(z|s)+3\bar{\theta}} \\
(v_2^*(s))^2-(2E(z|s)-3\bar{\theta})v_2^*(s)+E(z^2|s)-3\bar{\theta}E(z|s) &= 0 \\
v_2^*(s) &= E(z|s)-\frac{3\bar{\theta}}{2} \pm \sqrt{\frac{9\bar{\theta}^2}{4}-Var(z|s)}
\end{aligned}
$$

The model yields two equilibria: one with a low outside offer, the left root, and another with a high outside offer, the right root. However, the left root is unstable. Indeed, by offering a slightly higher wage, outside firms would attract workers and make a larger profit. The right root is stable: deviation to the right implies a negative profit, while deviation to the left results in no chance to attract a worker. Thus, I only consider right roots thereafter.

$$
\begin{aligned}
v_2^*(s) &= E(z|s)-\frac{3\bar{\theta}}{2} + \sqrt{\frac{9\bar{\theta}^2}{4}-Var(z|s)} \\
v_2^*(s) = E(z|s)-b_2\ with\ b_2 &= \frac{3\bar{\theta}}{2} - \sqrt{\frac{9\bar{\theta}^2}{4}-Var(z|s)}
\end{aligned}
$$

Expected value at $t = 2$

$$\mathrm{E}[\mathrm{V}_2|\tilde{z},s] = E[(z-w_2)P(\theta_2 < w_2 - v_2)|\tilde{z},s]$$
$$E[V_2|\tilde{z},s] = E[(z - \tfrac{z+v_2-\bar{\theta}}{2})(\tfrac{w_2-v_2+\bar{\theta}}{2\bar{\theta}})|\tilde{z},s]$$
$$E[V_2|\tilde{z},s] = E[(z - \tfrac{z+v_2-\bar{\theta}}{2})(\tfrac{\frac{z+v_2-\bar{\theta}}{2}-v_2+\bar{\theta}}{2\bar{\theta}})|\tilde{z},s]$$
$$E[V_2|\tilde{z},s] = E[(z-v_2+\bar{\theta})^2|\tilde{z},s].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = E[(z-E(z|s)+b_2+\bar{\theta})^2|\tilde{z},s].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = E[(z-E(z|s))^2+(b_2+\bar{\theta})^2+2(z-E(z|s))(b_2+\bar{\theta})|\tilde{z},s].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = E[z^2+E(z|s)^2-2zE(z|s)+(b_2+\bar{\theta})^2+2(z-E(z|s))(b_2+\bar{\theta}))|\tilde{z},s].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = [E(z^2|\tilde{z},s)+E(z|s)^2-2E(z|\tilde{z},s)E(z|s)+E(z|\tilde{z},s)^2-E(z|\tilde{z},s)^2+ (b_2+\bar{\theta})^2+2(E(z|\tilde{z},s)-E(z|s))(b_2+\bar{\theta})].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = [E(z|\tilde{z},s)^2+E(z|s)^2-2E(z|\tilde{z},s)E(z|s)+(b_2+\bar{\theta})^2+2(E(z|\tilde{z},s)- E(z|s))(b_2+\bar{\theta})E(z^2|\tilde{z},s)-E(z|\tilde{z},s)^2].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = [(E(z|\tilde{z},s)-E(z|s))^2+(b_2+\bar{\theta})^2+2(E(z|\tilde{z},s)-E(z|s))(b_2+\bar{\theta})+ E(z^2|\tilde{z},s)-E(z|\tilde{z},s)^2].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = [(E(z|\tilde{z},s)-E(z|s)+b_2+\bar{\theta})^2+V(z|\tilde{z},s)].\tfrac{1}{8\bar{\theta}}$$
$$E[V_2|\tilde{z},s] = \tfrac{1}{8\bar{\theta}}[\overbrace{(E(z|\tilde{z},s)-E(z|s)}^{\textbf{(Rent)}} + \underbrace{b_2+\bar{\theta}}_{\textbf{(penalties)}})^2 + \overbrace{Var(z|\tilde{z},s)}^{\textbf{(precision)}}]$$

Optimal Outside Option at t=1, v_1^*

$$\mathrm{v}_1^*(s) = E(z|s, worker\ rejects\ w_1^*)$$
$$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(1-P_{accepts}(1))\pi(z|s)}{\int_{-\infty}^{\infty}(1-P_{accepts}(1))\pi(z|s)}$$
$$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(1-\frac{w_1-v_1+\bar{\theta}}{2\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} 1-\frac{w_1-v_1+\bar{\theta}}{2\bar{\theta}}\pi(z|s)}$$
$$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(1-\frac{\frac{\tilde{y}_1+E(V_2|\tilde{z},s)+v_1-\bar{\theta}}{2}-v_1+\bar{\theta}}{2\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} 1-\frac{\frac{\tilde{y}_1+E(V_2|\tilde{z},s)+v_1-\bar{\theta}}{2}-v_1+\bar{\theta}}{2\bar{\theta}}\pi(z|s)}$$

$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(\frac{v_1 - \tilde{y}_1 + E(V_2|\tilde{z},s) + 3\bar{\theta}}{4\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} \frac{v_1 - \tilde{y}_1 + E(V_2|\tilde{z},s) + 3\bar{\theta}}{4\bar{\theta}}\pi(z|s)}$

$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(v_1 - \tilde{y}_1 + E(V_2|\tilde{z},s) + 3\bar{\theta})\pi(z|s)}{\int_{-\infty}^{\infty} (v_1 - \tilde{y}_1 + E(V_2|\tilde{z},s) + 3\bar{\theta})\pi(z|s)}$

$Numerator\ = E(z|s)v_1 - E(z\tilde{y}_1|s) - E(zE(V_2|\tilde{y}_1, s)|s) + 3\bar{\theta}E(z|s)$

$E(z\tilde{y}_1|s) = E(z|s)E(E(z|\tilde{y}_1, s)|s) + cov(z, E(z|\tilde{y}_1, s)|s)$

$E(z\tilde{y}_1|s) = E(z|s)^2 + Var(z|s) = E(z^2|s)$

$E(zE(V_2|\tilde{y}_1, s)|s) = E(z|s)E(E(V_2|\tilde{y}_1, s)|s) + cov(z, E(V_2|\tilde{y}_1, s)|s)$

$E(E(V_2|\tilde{y}_1, s)|s) = \frac{1}{8\bar{\theta}}.E([(E(z|\tilde{z}, s) - E(z|s) + b_2 + \bar{\theta})^2 + Var(z|\tilde{z}, s)]|s)$

$E(E(V_2|\tilde{y}_1, s)|s) = \frac{1}{8\bar{\theta}}.E([E(z|\tilde{z}, s)^2 + E(z|s)^2 - 2E(z|\tilde{z}, s)E(z|s) + (b_2 + \bar{\theta})^2 +$

$2(b_2 + \bar{\theta})(E(z|\tilde{z}, s) - E(z|s)) + Var(z|\tilde{z}, s)]|s)$

$E(E(z|\tilde{z}, s)^2|s) = E(E(z|\tilde{z}, s)|s)E(E(z|\tilde{z}, s)|s) + Var(E(z|\tilde{z}, s)|s)$

$E(E(z|\tilde{z}, s)^2|s) = E(z|s)^2 + Var(E(z|\tilde{z}, s)|s)$

$E(E(z|s)^2) = E(z|s)^2 - 2E(E(z|\tilde{z}, s)E(z|s)) = -2(E(E(z|\tilde{z}, s)|s).E(E(z|s)|s) +$

$cov(E(z|\tilde{z}, s), E(z|s)|s)) - 2E(E(z|\tilde{z}, s)E(z|s)) = -2(E(z|s)^2 + cov(E(z|\tilde{z}, s), E(z|s)|s)) =$

$-2E(z|s)^2$

$So, E(E(V_2|\tilde{y}_1, s)|s) = \frac{1}{8\bar{\theta}}.[E(z|s)^2 + Var(E(z|\tilde{z}, s)|s) + E(z|s)^2 - 2E(z|s)^2 +$

$(b_2 + \bar{\theta})^2 + E(Var(z|\tilde{z}, s)|s)])$

$E(E(V_2|\tilde{y}_1, s)|s) = \frac{1}{8\bar{\theta}}.[E(z|s)^2 + Var(E(z|\tilde{z}, s)|s) + E(Var(z|\tilde{z}, s)|s) - E(z|s)^2 +$

$(b_2 + \bar{\theta})^2]$

The Conditional Variance Formula: Var(z | s) = Var(E(z|,s) | s) + E(Var(z|,s) | s)

So, E(E($V_2|\tilde{y}_1, s)|s) = \frac{1}{8\bar{\theta}}.[Var(z|s) + (b_2 + \bar{\theta})^2]$

Numerator $= E(z|s)v_1 - E(z^2|s) - \frac{1}{8\bar{\theta}}.E(z|s).[Var(z|s) + (b_2 + \bar{\theta})^2] + 3\bar{\theta}E(z|s)$

Denominator $= v_1 + 3\bar{\theta} - E(\tilde{y}_1|s) - E(E(V_2|\tilde{z}, s)|s)$

Denominator $= v_1 + 3\bar{\theta} - E(z|s) - \frac{1}{8\bar{\theta}}.[Var(z|s) + (b_2 + \bar{\theta})^2]$

$v_1 = \frac{E(z|s)v_1 - E(z^2|s) - \frac{1}{8\bar{\theta}}.E(z|s).[Var(z|s) + (b_2 + \bar{\theta})^2] + 3\bar{\theta}E(z|s)}{v_1 + 3\bar{\theta} - E(z|s) - \frac{1}{8\bar{\theta}}.[Var(z|s) + (b_2 + \bar{\theta})^2]}$

$v_1^2 + v_1(3\bar{\theta} - 2E(z|s) - \frac{1}{8\theta}.[Var(z|s) + (b_2+\bar{\theta})^2]) - E(z^2|s) + \frac{1}{8\theta}.E(z|s).[Var(z|s) + (b_2 + \bar{\theta})^2] - 3\bar{\theta}E(z|s) = 0$

$v_1^*(s) = E(z|s) - [\frac{3\bar{\theta}}{2} - \frac{1}{16\theta}(Var(z|s) + (b_2+\bar{\theta})^2)] \pm \sqrt{[\frac{3\bar{\theta}}{2} - \frac{1}{16\theta}(Var(z|s) + (b_2 + \bar{\theta})^2)]^2 - Var(z|s)}$

$v_1^*(s) = E(z|s) - b_1$

$b_1 = [\frac{3\bar{\theta}}{2} - \frac{1}{16\theta}(Var(z|s) + (b_2+\bar{\theta})^2)] - \sqrt{[\frac{3\bar{\theta}}{2} - \frac{1}{16\theta}(Var(z|s) + (b_2 + \bar{\theta})^2)]^2 - Var(z|s)}$

Recalling that $b_2 = \frac{3\bar{\theta}}{2} - \sqrt{(\frac{3\bar{\theta}}{2})^2 - Var(z|s)}$, $b_1 > b_2$ because the function $x - \sqrt{x^2 - a}$ is a decreasing function of x with "a" as a constant.

Optimal incumbent wage at t=1, w_1^*

$$w_1^* = \frac{\tilde{y}_1 + E(V_2|\tilde{z}, s) + v_1 - \bar{\theta}}{2}$$

$$w_1^* = \frac{E(z|\tilde{z}, s) + \frac{1}{8\theta}.[(E(z|\tilde{z}, s) - E(z|s) + b_2 + \bar{\theta})^2 + Var(z|\tilde{z}, s)] + v_1 - \bar{\theta}}{2}$$

$$w_1^* = \frac{E(z|\tilde{z}, s) + \frac{1}{8\theta}.[((E(z|\tilde{z}, s) - E(z|s))^2 + (b_2 + \bar{\theta})^2 + 2(b_2 + \bar{\theta})(E(z|\tilde{z}, s) - E(z|s)) +}{2}$$

$$+ \frac{Var(z|\tilde{z}, s)] + E(z|s) - b_1 - \bar{\theta}}{2}$$

$$w_1^* = \frac{E(z|\tilde{z}, s) + \frac{2(b_2+\bar{\theta})}{8\theta}E(z|\tilde{z}, s) + E(z|s) - \frac{2(b_2+\bar{\theta})}{8\theta}E(z|s) + \frac{1}{8\theta}.[(E(z|\tilde{z}, s) - E(z|s))^2 + (b_2 + \bar{\theta})^2]}{2}$$

$$- \frac{(b_1 + \bar{\theta})}{2}$$

$$w_1^* = \frac{\frac{(5+\frac{b_2}{\theta})}{4}E(z|\tilde{z}, s) + \frac{(3-\frac{b_2}{\theta})}{4}E(z|s) - (b_1 + \bar{\theta}) - \frac{1}{8\theta}.[(E(z|\tilde{z}, s) - E(z|s))^2 + (b_2 + \bar{\theta})^2]}{2}$$

$$w_1^* = \frac{\frac{(5+\frac{b_2}{\theta})}{4}E(z|\tilde{z}, s) + \frac{(3-\frac{b_2}{\theta})}{4}E(z|s) - (b_1 + \bar{\theta}) - \frac{1}{8\theta}.K}{2}$$

$$with\ K = (E(z|\tilde{z}, s) - E(z|s))^2 + (b_2 + \bar{\theta})^2]$$

Proof of Proposition 1

Proposition 1: *Employer learning and layoffs — Case of increasing asymmetry*

Holding schooling constant, the layoff hazard rate is negatively related to ability at both t= 1 and t=2, with a larger magnitude at time t=2.

Holding ability constant, the layoff hazard rate is positively related to schooling at both t=1 and t=2, with a larger magnitude at t=2.

$\lambda_L(1) = P[E(z|\tilde{z}, s) - w_1^* < 0]$

$\lambda_L(1) = P[\frac{1}{2}(\frac{1}{4}(3 - \frac{b_2}{\bar{\theta}})(E(z|\tilde{z}, s) - E(z|s)) + b_1 + \bar{\theta} - \frac{1}{8\bar{\theta}}K) < 0]$

$with\ E(z|\tilde{z}, s) = \beta\tilde{z} + \gamma s = \beta(z + \epsilon) + \gamma s$

$E(z|s) = \alpha s$

$\lambda_L(1) = P[\epsilon < \frac{4\bar{\theta}}{\beta(3\bar{\theta} - b_2)}[\frac{1}{8\bar{\theta}}K - (b_1 + \bar{\theta})] + \frac{\alpha - \gamma}{\beta}s - z]$

$\lambda_L(1) = F_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta} - b_2)}[\frac{1}{8\bar{\theta}}K - (b_1 + \bar{\theta})] + \frac{\alpha - \gamma}{\beta}s - z)$

$\frac{\partial \lambda_L(1)}{\partial z}_{|s} = -\int_{(z,s)} f_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta} - b_2)}[\frac{1}{8\bar{\theta}}K - (b_1 + \bar{\theta})] + \frac{\alpha - \gamma}{\beta}s - z)d(z, s) = -1 < 0$

$\frac{\partial \lambda_L(1)}{\partial s}_{|z} = \frac{\alpha - \gamma}{\beta}\int_{(z,s)} f_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta} - b_2)}[\frac{1}{8\bar{\theta}}K - (b_1 + \bar{\theta})] + \frac{\alpha - \gamma}{\beta}s - z)d(z, s) = \frac{\alpha - \gamma}{\beta} > 0$

$\alpha > \gamma$

$\lambda_L(2) = P[\frac{z - E(z|s) + b_2 + \bar{\theta}}{2} < 0|E(z|\tilde{z}, s) - w_1^* >= 0, w_1^* - \theta_1 >= v_1^*]$

$\lambda_L(2) = P[\frac{z - E(z|s) + b_2 + \bar{\theta}}{2} < 0|E(z|\tilde{z}, s) - w_1^* >= 0, w_1^* - \theta_1 >= v_1^*]$

$\lambda_L(2) = P[\frac{z - E(z|s) + b_2 + \bar{\theta}}{2} < 0]$

$\lambda_L(2) = 1 - F_s(\frac{z + b_2 + \bar{\theta}}{\alpha})$

$\frac{\partial \lambda_L(2)}{\partial z}_{|s} = -\frac{1}{\alpha}\int_z f_s(z + b_2 + \bar{\theta})dz = -\frac{1}{\alpha} < 0$

$abs(\frac{\partial \lambda_L(2)}{\partial z}_{|s}) > abs(\frac{\partial \lambda_L(1)}{\partial z}_{|s})$

$\alpha < 1$

$\lambda_L(2) = F_z(\alpha s - b_2 - \bar{\theta})$

$\frac{\partial \lambda_L(2)}{\partial s}_{|z} = \int_s f_z(\alpha s - b_2 - \bar{\theta}) ds = 1 > 0$

$abs(\frac{\partial \lambda_L(2)}{\partial s}_{|z}) > abs(\frac{\partial \lambda_L(1)}{\partial s}_{|z})$

$\frac{\alpha-\gamma}{\beta} < 1$

$\frac{\alpha-\gamma}{\beta} < 1\, if$ *is more correlated to z than s is to z, which should be the case if incumbent firms are learning.*

Proof of Proposition 2

Proposition 2: *Employer learning and quits — Case of increasing asymmetry*

Holding schooling constant, the quit hazard rate is negatively related to ability at both t=1 and t=2, with a larger impact at t=2. The magnitude of both effects depends on the distribution of θ. The more variability there is in θ, the lower both effects are.

Holding ability constant, the quit hazard rate is positively related to schooling at both t=1 and t=2 with a larger impact at t=2. The magnitude of both effects depends on the distribution of θ. The more variability there is in θ, the lower both effect are.

$\lambda_Q(1) = P[w_1^* - \theta_1 < v_1^* | E(z|\tilde{z}, s) - w_1^* >= 0]$

$\lambda_Q(1) = \frac{P[w_1^* - \theta_1 < v_1^* ; E(z|\tilde{z}, s) - w_1^* >= 0]}{P[E(z|\tilde{z}, s) - w_1^* >= 0]}$

$For x \in [-\bar{\theta}; \bar{\theta}]$

$\lambda_Q(1) = \frac{\int_x P[w_1^* < x + v_1^* ; \theta_1 = x ; E(z|\tilde{z}, s) - w_1^* >= 0] dx}{P[E(z|\tilde{z}, s) - w_1^* >= 0]}$

$\lambda_Q(1) = \frac{\int_x P[w_1^* < x + v_1^* ; E(z|\tilde{z}, s) - w_1^* >= 0] P[\theta_1 = x] dx}{P[E(z|\tilde{z}, s) - w_1^* >= 0]}$

$\lambda_Q(1) = \frac{\frac{1}{2\bar{\theta}} \int_x P[w_1^* < x + v_1^* ; E(z|\tilde{z}, s) - w_1^* >= 0] dx}{P[E(z|\tilde{z}, s) - w_1^* >= 0]}$

$w_1^* < x + v_1 \Leftrightarrow \epsilon < \frac{4\bar{\theta}}{\beta(5\bar{\theta} + b_2)}(2x + \bar{\theta} - b_1 - \frac{1}{8\bar{\theta}} K) + \frac{\alpha-\gamma}{\beta} s - z$

$E(z|\tilde{z}, s) - w_1^* \geq 0 \Leftrightarrow \epsilon \geq \frac{4\bar{\theta}}{\beta(3\bar{\theta} - b_2)}[\frac{1}{8\bar{\theta}} K - (b_1 + \bar{\theta})] + \frac{\alpha-\gamma}{\beta} s - z$

$Since x \in [-\bar{\theta};\bar{\theta}], 2x+\bar{\theta} \in [-\bar{\theta};3\bar{\theta}]$

$$\lambda_Q(1) = \frac{\frac{1}{2\bar{\theta}}P[\frac{4\bar{\theta}}{\beta(3\bar{\theta}-b_2)}[\frac{1}{8\bar{\theta}}K-(b_1+\bar{\theta})]+\frac{\alpha-\gamma}{\beta}s-z\leq\epsilon\leq\frac{4\bar{\theta}}{\beta(5\bar{\theta}+b_2)}(3\bar{\theta}-b_1-\frac{1}{8\bar{\theta}}K)+\frac{\alpha-\gamma}{\beta}s-z]}{1-F_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta}-b_2)}[\frac{1}{8\bar{\theta}}K-(b_1+\bar{\theta})]+\frac{\alpha-\gamma}{\beta}s-z)}$$

$$\lambda_Q(1) = \frac{1}{2\bar{\theta}}\frac{F_\epsilon(\frac{4\bar{\theta}}{\beta(5\bar{\theta}+b_2)}[3\bar{\theta}-b_1-\frac{1}{8\bar{\theta}}K]+\frac{\alpha-\gamma}{\beta}s-z)-F_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta}-b_2)}[\frac{1}{8\bar{\theta}}K-(b_1+\bar{\theta})]+\frac{\alpha-\gamma}{\beta}s-z)}{1-F_\epsilon(\frac{4\bar{\theta}}{\beta(3\bar{\theta}-b_2)}[\frac{1}{8\bar{\theta}}K-(b_1+\bar{\theta})]+\frac{\alpha-\gamma}{\beta}s-z)}$$

$$A = \frac{4\bar{\theta}}{\beta(5\bar{\theta}+b_2)}[3\bar{\theta}-b_1-\frac{1}{8\bar{\theta}}K]+\frac{\alpha-\gamma}{\beta}s-z$$

$$B = \frac{4\bar{\theta}}{\beta(3\bar{\theta}-b_2)}[\frac{1}{8\bar{\theta}}K-(b_1+\bar{\theta})]+\frac{\alpha-\gamma}{\beta}s-z$$

$$\frac{\partial\lambda_Q(1)}{\partial z}_{|s} = \frac{1}{2\bar{\theta}}\int_{(z,s)}\frac{[-f_\epsilon(A)+f_\epsilon(B)].(1-F_\epsilon(B))-f_\epsilon(B).[F_\epsilon(A)-F_\epsilon(B)]}{(1-F_\epsilon(B))^2}d(z,s) < 0$$

$$\frac{\partial\lambda_Q(1)}{\partial z}_{|s} = \frac{1}{2\bar{\theta}}\int_{(z,s)}\frac{[-f_\epsilon(A)+f_\epsilon(B)].(1-F_\epsilon(B))-f_\epsilon(B).[1-F_\epsilon(B)+F_\epsilon(A)-1]}{(1-F_\epsilon(B))^2}d(z,s) < 0$$

$$\frac{\partial\lambda_Q(1)}{\partial z}_{|s} = \frac{1}{2\bar{\theta}}\int_{(z,s)}\frac{-f_\epsilon(A).(1-F_\epsilon(B))+f_\epsilon(B).(1-F_\epsilon(A))}{(1-F_\epsilon(B))^2}d(z,s) < 0$$

$$\frac{\partial\lambda_Q(1)}{\partial z}_{|s} = \frac{1}{2\bar{\theta}}\int_{(z,s)}\frac{-f_\epsilon(A).(1-F_\epsilon(B))+f_\epsilon(B).(1-F_\epsilon(A))}{(1-F_\epsilon(B))^2}d(z,s) > -\frac{1}{2\bar{\theta}}$$

$$\frac{\partial\lambda_Q(1)}{\partial s}_{|z} = \frac{1}{2\bar{\theta}}\frac{\alpha-\gamma}{\beta}\int_{(z,s)}\frac{[f_\epsilon(A)-f_\epsilon(B)].(1-F_\epsilon(B))+\frac{\alpha-\gamma}{\beta}f_\epsilon(B).[F_\epsilon(A)-F_\epsilon(B)]}{(1-F_\epsilon(B))^2}d(z,s) > 0$$

$$\lambda_Q(2) = P[w_2^*-\theta_2 < v_2^*|E(z|\tilde{z},s)-w_1^* >= 0, w_1^*-\theta_1 >= v_1^*, z-w_2^* >= 0]$$

$$\lambda_Q(2) = 1-P[\theta_2 < \frac{z-E(z|s)+b_2-\bar{\theta}}{2}]$$

$$\lambda_Q(2) = 1-\frac{z-E(z|s)+b_2+\bar{\theta}}{2\bar{\theta}}$$

$$\frac{\partial\lambda_Q(2)}{\partial z}_{|s} = -\frac{1}{2\bar{\theta}} < 0$$

$$\frac{\partial\lambda_Q(2)}{\partial s}_{|z} = \frac{\alpha}{2\bar{\theta}} > 0$$

$$abs(\frac{\partial\lambda_Q(2)}{\partial z}_{|z}) > abs(\frac{\partial\lambda_Q(1)}{\partial z}_{|s})$$

$$abs(\frac{\partial\lambda_Q(2)}{\partial s}_{|z}) > abs(\frac{\partial\lambda_Q(1)}{\partial s}_{|z}) \text{ because } \frac{\alpha-\gamma}{\beta} < 1$$

.1.4 Case 2: Decreasing Asymmetry

In the case of decreasing asymmetry, the information set of outside firms widens and outside firms catch up with incumbent firms. Outside firms do not observe the signal of ability during the first period of employment, but they learn ability during the second period of employment.

Expected value at t=2

$$\mathrm{E}(\mathrm{V}_2|\tilde{z}, s) = E[(z - w_2)P(w_2 - \theta \geq v_2)|\tilde{z}, s]$$
$$E(V_2|\tilde{z}, s) = E[(z - z + \frac{\bar{\theta}}{2})P(\theta \leq \frac{\bar{\theta}}{2})]$$
$$E(V_2|\tilde{z}, s) = \frac{\bar{\theta}}{8}$$

Optimal Outside Option at t=1, v_1^*

$$\mathrm{v}_1^*(s) = E(z|s, worker\ rejects\ w_1^*)$$
$$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(1 - \frac{\frac{z+v_1-\frac{7\bar{\theta}}{8}}{2} - v_1 + \bar{\theta}}{2\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} 1 - \frac{\frac{z+v_1-\frac{7\bar{\theta}}{8}}{2} - v_1 + \bar{\theta}}{2\bar{\theta}}\pi(z|s)}$$
$$v_1^*(s) = \frac{\int_{-\infty}^{\infty} z(\frac{v_1 - z + \frac{23\bar{\theta}}{8}}{4\bar{\theta}})\pi(z|s)}{\int_{-\infty}^{\infty} \frac{v_1 - z + \frac{23\bar{\theta}}{8}}{4\bar{\theta}}\pi(z|s)}$$
$$v_1^*(s) = \frac{v_1^*(s)E(z|s) - E(z^2|s) + \frac{23\bar{\theta}}{8}E(z|s)}{v_1^*(s) - E(z|s) + \frac{23\bar{\theta}}{8}}$$
$$(v_1^*(s))^2 - (2E(z|s) - \frac{23\bar{\theta}}{8})v_1^*(s) + E(z^2|s) - \frac{23\bar{\theta}}{8}E(z|s) = 0$$
$$v_1^*(s) = E(z|s) - \frac{23\bar{\theta}}{16} \pm \sqrt{(\frac{23\bar{\theta}}{16})^2 - V(z|s)}$$
$$v_1^*(s) = E(z|s) - b$$
$$with\ b = \frac{23\bar{\theta}}{16} - \sqrt{(\frac{23\bar{\theta}}{16})^2 - V(z|s)}$$
$$Thus, b_2 < b < b_1.$$

Proof of Proposition 3

Proposition 3: *Employer learning and layoffs — Case of decreasing asymmetry*

Holding schooling constant, the layoff hazard rate is negatively related to ability at t=1. The model predicts no layoff at t=2.

Holding ability constant, the layoff hazard rate is positively related to schooling at t=1. The *model predicts no layoff at t=2.*

$\lambda_L(1) = P[E(z|\tilde{z}, s) \leq E(z|s) - b - \frac{7\bar{\theta}}{8}]$
$\lambda_L(1) = F_\epsilon(\frac{\alpha-\gamma}{\beta}s - \frac{8b+7\bar{\theta}}{8\beta} - z)$
$\frac{\partial\lambda_L(1)}{\partial z}_{|s} = -\int_{(z,s)} f_\epsilon(\frac{\alpha-\gamma}{\beta}s - \frac{8b+7\bar{\theta}}{8\beta} - z)d(z,s) = -1 < 0$
$\frac{\partial\lambda_L(1)}{\partial s}_{|z} = \frac{\alpha-\gamma}{\beta}\int_{(z,s)} f_\epsilon(\frac{\alpha-\gamma}{\beta}s - \frac{8b+7\bar{\theta}}{8\beta} - z)d(z,s) = \frac{\alpha-\gamma}{\beta} > 0$
No layoffs at t=2.

Proof of Proposition 4

Proposition 4: *Employer learning and Quits — Case of decreasing asymmetry*

Holding schooling constant, the quit hazard rate is negatively related to ability at t=1. Ability does not affect quits at t=2.

Holding ability constant, the quit hazard rate is positively related to schooling at t=1. Schooling does not affect quits at t=2.

$\lambda_Q(1) = P[\theta_1 \geq \frac{1}{2}(E(z|\tilde{z}, s) - E(z|s) + b - \frac{7\bar{\theta}}{8})|E(z|\tilde{z}, s) \geq E(z|s) - b - \frac{7\bar{\theta}}{8}]$
$\lambda_Q(1) = P[\epsilon \leq (\frac{1}{\beta}(2\theta_1 + (\alpha-\gamma) - b + \frac{7\bar{\theta}}{8}) - z|\epsilon \geq \frac{\alpha-\gamma}{\beta}s - \frac{8b+7\bar{\theta}}{8\beta} - z]$
$\lambda_Q(1) = \frac{\int_x P[\epsilon\leq(\frac{1}{\beta}(2x+(\alpha-\gamma)s-b+\frac{7\bar{\theta}}{8})-z,\theta_1=x,\epsilon\geq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]dx}{P[\epsilon\geq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]}$
$\lambda_Q(1) = \frac{1}{2\bar{\theta}}\frac{\int_x P[\epsilon\leq(\frac{1}{\beta}(2x+(\alpha-\gamma)s-b+\frac{7\bar{\theta}}{8})-z,\epsilon\geq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]dx}{P[\epsilon\geq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]}$
$\lambda_Q(1) = \frac{1}{2\bar{\theta}}\frac{\int_x P[\epsilon\leq(\frac{1}{\beta}(2x+(\alpha-\gamma)s-b+\frac{7\bar{\theta}}{8})-z]dx-P[\epsilon\leq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]}{1-P[\epsilon\leq\frac{\alpha-\gamma}{\beta}s-\frac{8b+7\bar{\theta}}{8\beta}-z]}$
$\int_x P[\epsilon \leq (\frac{1}{\beta}(2x + (\alpha-\gamma)s - b + \frac{7\bar{\theta}}{8}) - z]dx = [\frac{\beta}{2}((\frac{1}{\beta}(\frac{23\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z)F_\epsilon(\frac{1}{\beta}(\frac{23\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z) + f_\epsilon(\frac{1}{\beta}(\frac{23\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z)] - [\frac{\beta}{2}((\frac{1}{\beta}(\frac{-9\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z)F_\epsilon(\frac{1}{\beta}(-\frac{9\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z) + f_\epsilon(\frac{1}{\beta}(-\frac{9\bar{\theta}}{8} + (\alpha-\gamma)s - b) - z)]$

$A = \frac{1}{\beta}(\frac{23\bar{\theta}}{8} + (\alpha - \gamma)s - b) - z)$

$B = \frac{1}{\beta}(-\frac{9\bar{\theta}}{8} + (\alpha - \gamma)s - b) - z)$

$C = \frac{\alpha-\gamma}{\beta}s - \frac{8b+7\bar{\theta}}{8\beta} - z$

$\frac{\partial \lambda_Q(1)}{\partial z}_{|x} = \frac{1}{2\bar{\theta}} \int_{(z,x)} \frac{[-\frac{1}{2}(F_\epsilon(A)-F_\epsilon(B))+f_\epsilon(C)].(1-F_\epsilon(C)-f_\epsilon(C).[\int_x F_\epsilon((\frac{1}{\beta}(2x+(\alpha-\gamma)s-b+\frac{7\bar{\theta}}{8})-z)dx-F_\epsilon(C)]}{[1-F_\epsilon(C)]^2} d(z,s)$

$-\frac{1}{2\bar{\theta}} \leq \frac{\partial \lambda_Q(1)}{\partial z}_{|s} \leq 0$

$\frac{\partial \lambda_Q(1)}{\partial s}_{|z} = -\frac{\alpha-\gamma}{\beta} \frac{\partial \lambda_Q(1)}{\partial z}_{|s} > 0$

So that, $0 \leq \frac{\partial \lambda_Q(1)}{\partial z}_{|s} \leq \frac{1}{2\bar{\theta}}$

$\lambda_Q(2) = P[\theta_2 > -\frac{\bar{\theta}}{2} | E(z|\tilde{z}, s) \geq E(z|s) - b - \frac{7\bar{\theta}}{8}, \theta_1 \leq \frac{1}{2}(E(z|\tilde{z}, s) - E(z|s) + b - \frac{7\bar{\theta}}{8}]$

$\lambda_Q(2) = P[\theta_2 > -\frac{\bar{\theta}}{2}]$

$\frac{\partial \lambda_Q(2)}{\partial z}_{|s} = 0$

$\frac{\partial \lambda_Q(2)}{\partial s}_{|z} = 0$

.2 Appendix to Chapter 2

The solution of the consumer program verifies the following conditions:

For $i < i_w$:

$$c_{i,t} = \frac{e_t}{(1 + \tau_c)}, l_t = 0 \tag{13}$$

For $i >= i_w$:

$$u_{c_{i,t}} = \frac{\beta \phi_i u_{c_{i+1,t+1}}}{(1 + \gamma_r) q_t} \tag{14}$$

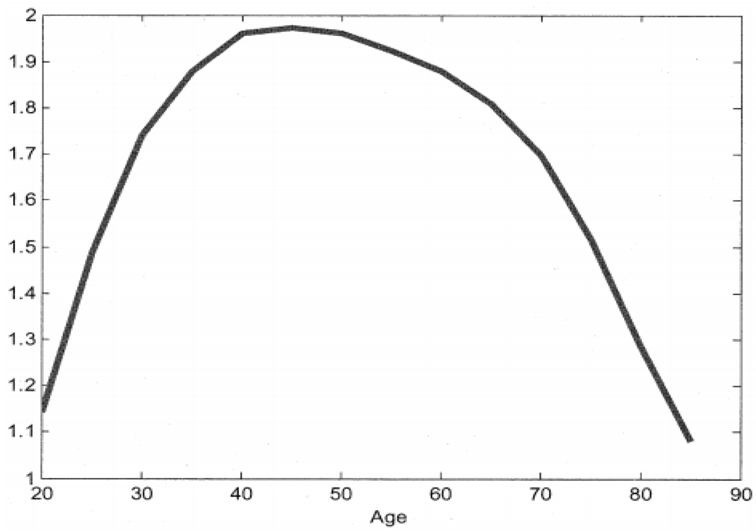


Figure 7: Age-profile of efficiency units of labor from Hansen (1993)

1. For $i < i_r$:

$$-\frac{u_{l_{i,t}}}{u_{c_{i,t}}} = \frac{(1-\tau_L)(1-\tau_p)\epsilon_i w_t}{1+\tau_C} \tag{15}$$

2. For $i >= i_r$:

$$l_{i,t} = 0 \tag{16}$$

Table 3: Death Probability by age

Cohort	Probability of death
20-24	0.00230
25-29	0.00195
30-34	0.00227
35-39	0.00310
40-44	0.00481
45-49	0.00749
50-54	0.01162
55-59	0.01837
60-64	0.02745
65-69	0.03949
70-74	0.05652
75-79	0.08172
80-84	0.11867
85-89	0.17086