The Inefficient Trickle-Down of Unemployment*

Regis Barnichon Yanos Zylberberg

JOB MARKET PAPER

Abstract

A substantial fraction of job holders are under-employed, i.e., in jobs for which they are over-qualified. This paper proposes a model to analyze the mechanisms behind under-employment and its consequences for the functioning of the labor market. We show that under-employment is generally inefficient: in a competitive market, there are too many high-skill workers employed in low-productivity firms, and there are not enough high productivity firms. Under-employment generates a trickle-down phenomenon, in which unemployment trickles down from the high-skill groups to the low-skill groups. The trickle-down of unemployment exacerbates inequality across worker-skill groups by redistributing shocks from the high-skill groups to the low skill groups. As a result, high-skill workers enjoy not only higher expected income but also lower income volatility.

JEL: E24, J22, J64.

Keywords: unemployment, directed search, misallocation.

*Barnichon: CREI, Universitat Pompeu Fabra; rbarnichon@crei.cat; Zylberberg: CREI, Universitat Pompeu Fabra; yzylberberg@crei.cat; Ramon Trias Fargas 25-27, 08005-Barcelona, Spain.
We thank Paula Bustos, Jan Eeckhout, Jordi Gali, Nicola Gennaioli, Alberto Martin, Jaume Ventura and seminar participants at CREI and Universitat Pompeu Fabra for very useful remarks and encouragements. Financial support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement nº 241114 is gratefully acknowledged by Yanos Zylberberg.

The recession left millions of college-educated working in coffee shops and retail stores.[1]

When choosing between different applicants, one should not hire the candidate with the highest education level. [...] The education level of the successful candidate should not exceed two years after high-school.[2]

While the recent crisis left many workers without a job, it also left many workers under-employed, i.e., in jobs for which they are overqualified. For instance, the fraction of US college graduates working in lower skill-requirement occupations was over 35 percent by the end of 2012 (figure 1). This large number reflects two stylized facts about under-employment. First, the labor market is constantly characterized by a large rate of under-employment and this rate has steadily increased over the past 30 years (figure 2). Second, the rate of under-employment is strongly-countercylical, mimicking the evolution of the unemployment rate (figure 3), and the under-employment rate increased strongly during the last recession.

Under-employment finds a large echo in the media and is often taken as *prima facie* evidence that labor markets are "malfunctioning". But is under-employment inefficient? And if so, what mechanism lies behind any possible inefficiency? To address these questions, we need to understand how an individual's decision to move down the occupational ladder can affect the labor market of other workers, and this paper provides one such theory.

We present a model of "trickle-down unemployment" in which some high-skill workers move-down the occupational ladder in order to escape competition from their high-skill peers and find a job more easily. In doing so, they take the jobs of less-skill individuals, who are in turn driven out their market and down the occupational ladder. We show that under-employment and the trickle-down of unemployment are generally inefficient. In the decentralized allocation, too many high-skill workers are under-employed, and there are too many low productivity firms and not enough high productivity firms.

A concrete example helps clarify the mechanism behind under-employment and the inefficient trickle-down of unemployment. Consider the job search problem of a young individual –call him Albert–, who just graduated from college in the middle of a recession. On the one hand, Albert can concentrate his job search on high-productivity firms with college degree requirements. However, jobs are scarce and

[1]The Wall Street Journal, 26 March 2013

[2]Hiring instructions for support positions in public schools under the "Jobs for the Young" program. Bulletin Officiel de l'Education Nationale, 16 December 1997

competition is fierce. Moreover, even if he gets an offer, his bargaining power will be low, because the firm has many other good candidates that could do the job. On the other hand, Albert can apply to middle-productivity firms with lower education requirements, such as a community college degree. Competition is also fierce (jobs are scarce everywhere), and he will not be the only applicant. However, this time, Albert will be considered an exceptional candidate and be very likely to get the job. Being one above many, his bargaining power will be relatively high. However, when Albert moves down the job ladder, he takes the job of a young individual, Bob, who just graduated from community college. As a result, Bob may decide to apply to low-productivity firms with only high-school degree requirements, which in turn will push some high-school graduates down the ladder. As high-school dropouts have nowhere to go, they will end up unemployed in greater proportion or will drop out of the labor force. In this chain of event, unemployment of the high-education group trickles down to the lower-education groups. This mechanism is generally inefficient, because when deciding to search for a job in a middle-productivity firm, Albert does not take into account how he disproportionately hurts the labor market opportunities of lower-skill workers.

The key ingredients of our model are heterogeneity across workers and jobs, coordination frictions, and crucially, wage competition between workers. Workers differ in their skill level, and islands differ in their productivity level. Workers can direct their search to a given island, that can be more or less productive and more or less congested. In each island, there are coordination frictions: some vacancies will receive multiple applications, while other vacancies will have no applicants, and not every worker will get a job. When a vacancy receives multiple application, the firm makes applicants compete for the job and hires the most profitable applicant. Hiring is *not random*. At the root of under-employment in our model lies a trade-off between the output of a job in that island and the level of competition for that job. In productive but congested islands, output is high but workers' bargaining position is low, because workers compete against each other and leave most of the surplus to the firm. In a less productive island, production is lower, but *even when congestion is high*, a high-skill worker is ranked above the other applicants and can easily get a job, because he is an exceptional candidate, being "better" than most other workers in that island.

Two stylized facts support the notion, central to our model, that high-skill workers are "better" and command a wage premium over their lower-skill peers, because they are ranked higher by firms. First, figure 4 shows that while college graduates suffer a wage loss when they become under-employed, they still earn a premium over non-college graduates working in the same occupations. Naturally, the larger the number of high-skill workers searching in a lower-tech island, the less “unique” they are, and the lower their wage premium over lower-skilled workers. Consistent with this prediction, figure 5 shows a strong negative correlation between under-employment and the wage premium paid to under-employed college graduates over non-college graduates working in the same occupations. While this negative correlation can be easily explained with non-random hiring and wage competition between workers, it is difficult to explain in a model where hiring is random and firms cannot choose among different applicants.[3]

We show that under-employment is generally inefficient. Although the optimal allocation calls for some level of under-employment in order to maximize the matching probability of the most-skilled groups, the market is characterized by a *ranking externality*, and there is too much under-employment in the decentralized allocation. The key element behind the ranking externality is the fact that high-skill workers and low-skill workers can face *different* labor markets in the *same* island, because a high-skill worker is ranked first by firms and systematically gets the job when faced with lower-skill applicants. As a result, the gain (for a high-skill worker) from searching in a low-productivity island can be strongly disconnected from the loss inflicted on the other workers in that island, and the inefficiency associated with the ranking externality can be large.[4] The ranking externality is strongest when the marginal high-skill applicant is most different from the average applicant, i.e., most “unique”.

To get some intuition about the ranking externality and how the presence of a high-skill worker can disproportionately hurt the labor market of low-skill workers, it is helpful to go back to an illustrative example with Albert –the high-skill worker– and Bob –the low-skill worker–. Imagine that the high-productivity island is little congested with many jobs, but that the low-productivity island is very congested with few jobs and many low-skill workers (including Bob). Imagine that there are yet no high-skill workers in the low-productivity island. In that case, the low-productivity island looks completely different to Bob and Alfred: the market is very congested for Bob, but it is very attractive for Alfred, because he faces no competition. As a result, Alfred may decide to *take* a job in the low-productivity island for a very small gain (Alfred’s market was little congested to begin with)

[3]As is the case in standard search models, such as random search and matching models (Mortensen and Pissarides (1994)) or directed search models (Moen (1997)).

[4]This implication of our model is in sharp contrast to models with random hiring (Mortensen and Pissarides (1994)): in those models, high-skill and low-skill always face the same labor market prospects in the same island. As a result, the gain from moving down the ladder for a high-skill worker is strongly related to the loss inflicted on the other workers in that island.

while imposing a substantial cost to Bob an his peers (Bob's labor market was very congested to begin with).

In the decentralized equilibrium, we show that, compared to the constrained optimal allocation (subject to the same coordination frictions), too many high-skill workers are "wasted" working for low productivity firms, too many low-skill workers end up unemployed, and there are too many low productivity firms and too few high productivity firms.

Under-employment has also interesting distributional implications. A higher skill level may not only guarantee a higher expected income, but it may also provide a lower volatility of income, because high-skill workers can partially smooth out adverse labor demand shocks by moving down the occupational ladder. In contrast, under-employment exacerbates the income volatility faced by lower-skill workers. For instance, while an adverse labor demand shock hitting low-productivity firms does not affect the expected income of high-skill workers, the converse is not true: an adverse labor demand shock hitting high-productivity firms always affects low-skill workers.

In a final section, we perform a simple empirical exercise to quantify the magnitude of the inefficiency associated with the ranking externality, as well as test two key predictions of the model: (i) under-employment generates downward pressures on wages in low-productivity firms, but (ii) under-employment does not affect wages in high-productivity firms. We find strong support for these predictions and show that the inefficiency associated with under-employment is likely to be large: a back-of-the-envelope calculation shows that reducing the extent of under-employment by 15 percent would increase total output by about 5 percent.

Studies of the phenomenon of under-employment, also referred to as over-education, goes back to the 1970s when the supply of educated workers seemed to outpace its demand in the labor market, apparently resulting in a substantial reduction in the returns to schooling.[5] While motivated with heterogeneity across education levels, the mechanisms and inefficiency we identify in this paper are more general and apply whenever workers display heterogeneity in productivity, be it determined by education or other factors.

Our approach to modeling under-employment builds on the competitive search literature (Moen 1997) –also known as directed search with wage posting–, in which firms post wage offers and workers can direct their search to the markets with the most attractive alternatives. Closely related are Shi (2001, 2002), Shimer (2005)

[5] See Freeman (1976), Thurow (1975), Sicherman (1991), Sicherman and Galor (1990) and Leuven and Oosterbeek (2011) for a review.

and Eeckhout and Kircher (2010) who study competitive search models with heterogeneous agents.[6] In contrast to that literature, we relax the assumptions of (i) random hiring, and (ii) perfect wage commitment. The possibility for the firm to rank applicants, coupled with the absence of wage commitment, is at the heart of the inefficiency generated by under-employment. This result contrasts with competitive search models, in which the decentralized allocation is efficient.

Our modeling of non-random hiring, in which firms can choose among different applicants, builds on Blanchard and Diamond (1994) idea of ranking, in which firms rank applicants according to some observed criteria. In our model, the productivity of each applicant is observed, and firms can rank applicants according to their productivity level.

With the absence of a commitment technology, firms cannot commit to the posted terms of trade, because firms and workers realize that, upon forming a match, either party will want to renegotiate over the posted wage.[7] A contribution of this paper is to propose a tractable and intuitive bargaining setup that can capture wage negotiations with (i) multiple, and (ii) heterogeneous, applicants in a non-random hiring setting. Our modeling of wage negotiation is related to the competing-auction theories of Shimer (1999) and Julien et al. (2000), in which job candidates auction their labor services to employers. Wage bargaining departs from the standard Nash-bargaining outcome (e.g., Pissarides (2000)), because firms can collect applications and make applicants compete for the job. As a result, the outside option of the firm, and thus the surplus extracted by the firm depends on the number and on the type (or quality) of the other applicants.

The possibility that unemployment may trickle-down to lower layers recently received empirical support in A. Gautier et al. (2002) and Beaudry et al. (2013). In particular, Beaudry et al. argue that, around the year 2000, the demand for skill underwent a reversal, which led high-skill workers to move down the occupational ladder and push low-skilled workers even further down the occupational ladder.

The remainder of this paper is structured as follows. In section 1, we describe the environment and the wage negotiation process. In section 2, we characterize the partial equilibrium allocation with exogenous labor demand, first with 2 islands and 2 types of workers, then with 3 islands and 3 types of workers in order to study propagation between islands. Section 3 introduces endogenous firm entry and studies how the partial equilibrium allocation is modified by firms' reaction to workers

[6] See also Albrecht and Vroman (2002) and Gautier (2002) for models with worker and firm heterogeneity but with random search and Nash bargaining.

[7] We are not the first ones to relax this assumption. See for instance, Coles (2001) and, more recently Doyle and Wong (2013) in models of standard search and matching.

movements. Section 4 discusses the optimality of the decentralized allocation. Finally, section 5 presents a simple test of the model's predictions and quantifies the magnitude of the inefficiency. We briefly conclude in section 6.

1 The model

This section presents the structure of the model and describes its three key ingredients: (i) heterogeneity across workers and jobs, (ii) search frictions, and (iii) wage competition between workers.

1.1 Preferences, technology and market structure

The model is static and consists of one period. There are two types of risk neutral agents in the economy, workers and firms, and the economy is divided into N islands indexed by $n = 1 \dots N$.

Firms operating in island n are characterized by a technology level ϕ_n. Islands are indexed such that $\phi_1 < \ldots < \phi_N$, i.e. island N has the highest technology level. A firm consists of one vacancy, and a firm can enter an island n by posting a vacancy at a cost $c_n > 0$. The number of firms/vacancies in each island will be determined endogenously by firm entry.

Workers are divided into N different types characterized by different productivity levels $p_1 < \ldots < p_N$. There is a mass L_n of agents of type n. A worker with a job provides inelastically one unit of labor to the firm and receives a salary ω. A worker without a job receives 0.[8] There is complementarity between workers' skill and firms' technology: a firm with technology ϕ_i paired with a worker with productivity p_j produces $\phi_i \times p_j$ for any $i, j \in \{1, \ldots, N\}^2$.

1.2 Coordination frictions

It takes time to match workers with jobs. To capture this search friction, we assume that each worker must settle to one island and can apply to at most one job. In a large anonymous market, workers cannot coordinate on which firm to apply to, leading to coordination frictions in each island. Some firms will get multiple applications, while others receive none. Some firms will receive applications from workers of different types, while others will receive applications from workers of the same type. Unmatched jobs and workers produce nothing and get 0 payoff.

With workers applying at random in a market with many workers and firms, the matching process is described by an urn-ball matching function (Butters (1977)),

[8] This assumption could be relaxed and is used for analytical simplicity.

in which each application (ball) is randomly allocated to a vacancy (urn). With a large number V of vacancies and a large number L of homogeneous applicants, the probability for a firm to be matched with exactly a applicants follows a multinomial distribution which can be approximated with a Poisson distribution

$$P(a) = \frac{q^a}{a!}e^{-q}$$

with $q = L/V$ the queue length, i.e. the job candidate to vacancy ratio.

The probability that a firm has exactly one applicant is then $P(1) = qe^{-q}$ and the probability that a worker finds a firm with no other candidate is $\frac{V}{L}P(1) = e^{-q}$.

This urn-ball matching function can be easily generalized to heterogeneous applicants. For instance, when there are two types of applicants, say, L_1 of type 1 and L_2 of type 2, we can proceed in a similar way and define the probability $P(a_1, a_2)$ that a firm faces a_1 applicants of types 1 and a_2 of type 2. Defining $q_1 = L_1/V$ and $q_2 = L_2/V$, we have

$$P(a_1, a_2) = \frac{q_1^{a_1}}{a_1!}e^{-q_1}\frac{q_2^{a_2}}{a_2!}e^{-q_2}.$$

In this setup, we can derive a number of probabilities that will be useful later. The probability that a worker type 1 is the only applicant to a job is $\frac{V}{L_1}P(a_1 = 1, a_2 = 0) = \frac{V}{L_1}q_1e^{-q_1-q_2} = e^{-q_1-q_2}$. Similarly, the probability that a worker type 1 is the only applicant of his type but with applicants of the other type is $e^{-q_1}(1-e^{-q_2})$. Finally, the probability that a worker type 1 faces other applicants of his type is $1 - e^{-q_1}$.

The general formula with N types follows easily from a similar reasoning. The probability $P(a_1, \ldots, a_N)$ that a firm faces a vector $(a_1, \ldots, a_N)$ of applicants is

$$P(a_1, \ldots, a_N) = \frac{q_1^{a_1}}{a_1!}e^{-q_1}\ldots\frac{q_1^{a_N}}{a_N!}e^{-q_N}.$$

1.3 Wage negotiation and hiring decision

In this section, we describe the wage bargaining process taking place between a firm and its (possibly multiple) job candidate(s). Specifically, we present a tractable and intuitive bargaining setup that can capture wage negotiations with (i) multiple and (ii) heterogeneous applicants in a non-random hiring setting.

A key dimension captured by the wage negotiation process is that, because workers compete against each other to get a job, the bargaining position of an applicant is endogenous and depends on (i) the tightness of the market (capturing the degree of competition a worker faces) and (ii) the quality of the unemployment pool (capturing the type of competitors an applicant faces).

There is perfect information, and all agents observe the pool of applicants and their types. We posit that the firm can only negotiate with one applicant at a time. Since the surplus of a match is highest with the most productive applicant, the firm chooses its best applicant and engages in an infinite horizon game of alternating offers as in Rubinstein (1982) in which the applicant announces his offer or announces if he accepts or rejects the firm's offer. If there are more than one applicant of the best type, the firm picks one at random.

Before negotiations with the chosen applicant starts, we allow the firm to sign with another applicant a "contingency" work contract, that would come into effect in case the initial negotiations break down. Since the outside option of the second-best applicant is to be unemployed and receive 0, we assume that the second-best applicant commits to work for a share ϵ, arbitrary small, of the surplus. The outside option of the firm is thus the surplus of hiring the second-best applicant.[9]

There are then four cases that determine the negotiated wage:

1. The firm has only one applicant. The bargaining game between that applicant and the firm is identical to that of Rubinstein (1982) and the worker receives a share β of the surplus, where β is determined by the relative impatience of the worker and the firm, or by the relative degrees of risk aversion as in Binmore et al. (1986):

$$\omega = \beta\phi p.$$

2. The firm has more than one applicants of *identical types*. Since the best applicant and the second-best applicant are identical, the firm gets all the surplus from the match, and the best applicant gets its reservation wage:[10]

$$\omega = 0.$$

[9]This assumption is made for simplicity as it provides a very tractable wage bargaining rule. Allowing for a more general negotiation wage rule would complicate the analysis but would not affect the crucial property that the outside option of the firm depends on the quality and number of the other applicants. For instance, one could think of a more general framework in which the firm starts a game of alternating offers with the best applicant, and that, if these negotiations break-down, the firm starts a game of alternating offers with the second-best applicant, then with the third-best if these negotiations break-down, etc.. While more cumbersome, the outcome of the negotiation would be qualitatively similar: the best-applicant would get the job and his wage would depend negatively on the productivity and number of other applicants.

[10]The outcome of this bargaining game can be also seen as workers bidding themselves down to their reservation wage. Note that because the firm gets all of the surplus, in this case, the worker is indifferent between employment and unemployment

3. If the firm has more than one applicants of *different types*. There are two subcases:

   (a) The firm has more than one applicants of the higher-type (for any number of lower-type applicants). The game is identical to case 2, and the firm gets all the surplus: $\omega = 0$.

   (b) The firm has one applicant of the higher-type and (one or many) lower-type applicants. Denote p_2 the productivity of the high-type and $p_1 < p_2$ the productivity of the second-best applicant.

   The firm and the high-type applicant engage in an infinite horizon game of alternating offers as in case 1, but for one crucial difference: The outside option of the firm is to hire the second-best applicant and get all the surplus, i.e., $p_1 \times \phi$. As a result, the worker gets a share β of the surplus generated over hiring the second-best applicant, i.e.,

$$\omega = \beta\phi(p_2 - p_1)$$

1.4 Timing

The timing of events is as follows. (1) Each worker chooses which island to send an application to. In parallel, each potential firm entrant decides whether to post a vacancy in any given island; (2) In each island, applications are randomly allocated to vacancies; (3) A wage negotiation ensues between the firm and its (possibly multiple) applicants; (4) Firm-worker matches are formed and production starts. Firms pay workers and realize profits.

To summarize, the firm's share of the surplus depends on the number *and* quality of applicants. If the firm can make applicants compete for the job, it is going to extract most of the surplus. However, it will not systematically capture all of the surplus because of worker heterogeneity. If one (and only one) applicant is strictly better (i.e., more productive) than the others, that applicant can extract some of the surplus, because the firm has a strict preference for that candidate and *ranks* him higher.

2 Equilibrium with Exogenous Labor Demand

In order to first clarify how workers decide on which island to search, this section characterizes the partial equilibrium with exogenous labor demand, i.e., taking the number of firms/vacancies in each island as given.

Definition 1. *Partial Equilibrium Allocation.*

Each worker of each type $i \in \{1, \ldots, N\}$ *decides on which island* $n \in \{1, \ldots, N\}$ *to search for a job, and the equilibrium is given by a sequence of worker choices* $C_i : [0, L_i] \mapsto \{1, \ldots, N\}$*, a matching function between applicants and firms and a sequence of allocations and wages.*

We first describe the equilibrium allocation with 2 islands and worker types, and then consider the more general case with 3 islands and worker types in order to study the propagation of under-employment across islands. Since we consider equilibria with under-employment, the Appendix describes the conditions that ensure the existence of a strictly positive rate of under-employment. Crucially however, the mechanisms we identify in this paper are not specific to under-employment, but arise whenever an island is, in equilibrium, populated by workers of heterogeneous productivity levels.

2.1 Some useful notations

We first introduce some useful notations. Let $h_n = \frac{L_n}{L_{n-1}}$ denote the size of the pool of type n workers relative to that of type $n-1$. For instance, a low h_n captures a situation in which the distribution of worker types is pyramidal: there are few high types and many low types. This means that high types will not crowd each other out (i.e., compete against each other), when they move down the occupation ladder. In other words, h_n captures how type n workers can "dilute" themselves in the islands below.

Let $q_n = \frac{L_n}{V_n}$ denote the ratio of type n individuals to job openings in island n. q_n can be seen as an "initial" queue length in island n, i.e., an hypothetical queue length corresponding to the case where workers of type n were assigned to island n and not allowed to move.

Denote x_n the fraction of workers in island n who move to island $n-1$ to look for a job. Starting from an "initial" number L_n of job candidates in island n, the number of workers type n who end up searching in island n is $(1 - x_n)L_n$, and the number of workers type n who search in an island below is $x_n L_n$. In island $n-1$, the ratio of searchers from island n to vacancies is thus given by $h_n x_n q_{n-1}$. In island n, the ratio of searchers from island n to vacancies is given by $q_n (1 - x_n)$.

Finally, denote $E\omega_{n,m}$ the expected income of an individual of type n searching in island m.

2.2 Equilibrium with $N = 2$: two worker types and two firm types

This section characterizes the equilibrium allocation and presents some comparative statics exercises to illustrate the mechanisms underlying the equilibrium.

Equilibrium In equilibrium, workers allocate themselves across islands until the point where the expected income in each island are equalized. Specifically, the equilibrium with two types of workers and firms is characterized by the following Proposition:

Proposition 1. *With $N = 2$, there is a unique equilibrium allocation of workers satisfying*

- *Type 2 workers are indifferent between islands 1 and 2, and x_2, the share of type 2 workers searching in island 1, is given by the arbitrage condition*

$$A(x_2) = -E\omega_{2,2} + E\omega_{2,1} = 0$$

with

$$\begin{cases} E\omega_{2,2} = \beta e^{-q_2(1-x_2)} p_2 \phi_2 \\ E\omega_{2,1} = \beta e^{-q_1 x_2 h_2} e^{-q_1} p_2 \phi_1 + \beta e^{-q_1 x_2 h_2} \phi_1 (p_2 - p_1) \left[1 - e^{-q_1}\right] \end{cases}$$

- *Type 1 workers only look for jobs in island 1 and their expected income is*

$$E\omega_{1,1} = \beta e^{-q_1(1+x_2 h_2)} p_1 \phi_1$$

Proof. Appendix. □

Each worker searches for a job in the island that provides him with the highest expected wage: in equilibrium, a type 2 worker is indifferent between looking for a job in island 2 and looking for a job in island 1, while a type 1 worker strictly prefers looking for a job in island 1. The arbitrage condition, $A(x_2)$, determines, x_2, the equilibrium allocation of type 2 workers across the two islands.

Figure 6 depicts the equilibrium allocation of type 2 workers as the intersection of the $E\omega_{2,1}$ curve, the expected wage earned in island 1, and the $E\omega_{2,2}$ curve, the expected wage earned in island 2. The $E\omega_{2,2}$ curve is increasing in x_2: an increase in the fraction of type 2 workers searching in island 1 lowers congestion in island 2, which lessens the competition type 2 workers face in island 2 and increases the expected wage. In contrast, an increase in the fraction of type 2 workers searching in island 1 makes island 1 more congested, which increases the competition workers face in island 1 and lowers the expected wage.

Comparative Statics We now present some comparative statics exercises to illustrate the mechanisms underlying the equilibrium allocation. In particular, we study how the under-employment rate depends on (i) market conditions in each island, (ii) the variance across workers' productivity, and (iii) the distribution of worker types.

As shown in Figure 7, an adverse labor demand shock affecting the high productivity island, i.e., an increase in the queue length q_2, shifts the expected wage earned in island 2, the $E\omega_{2,2}$ curve, to the right and thus generates a higher equilibrium x_2 and a lower expected wage for type 2 agents in both islands. Note that type 1 workers also see a decrease in their expected earnings (as shown by the $E\omega_{1,1}$ curve), because they face more competition from type 2 workers searching in island 1. In other words, a shock affecting the high type groups *trickles down* to the lower occupation group.[11]

As shown in Figure 8, a decrease in type 1 productivity, p_1, which increases the productivity difference between high-skill and low-skill workers, increases type 2 workers' expected wage in island 1, because with a lower p_1, a type 2 is more "unique". As a result, when type 2 workers compete with type 1 workers for a job, they extract more surplus from the firm. With an inward shift of the $E\omega_{2,1}$ curve, more type 2 workers move down to island 1, and their expected wage is higher. As a result, type 1 workers suffer (i) directly from their productivity loss (the shift in the $E\omega_{1,1}$ curve) and (ii) indirectly from the increased under-employment rate and increased competition of type 2 workers.

Finally, the distribution of worker types, the "shape of the pyramid", also affects the allocation and the extent of under-employment: a lower h_2, i.e., a pyramid with a larger base, allows type 2 workers to dilute themselves more easily in island 1. As a result, it leads to an outward shift of the $E\omega_{2,1}$ curve and thus to more under-employment.

2.3 Equilibrium with $N = 3$: three worker types and three firm types

The $N = 2$ case is a good benchmark to understand how workers decide on which island to search. However, it has no "propagation mechanism", in the sense that type 1 workers cannot respond to the competition of type 2 workers by moving

[11] A decline in the productivity differences between firms in islands 1 and 2 has a similar effect. Following a decline in ϕ_2, type 2 workers have less incentives to search in island 2, the $E\omega_{2,2}$ curve shifts to the right, and more type 2 workers move down to island 1.

further down the occupation ladder. To capture this possibility, we now study an economy with 3 islands and 3 worker types. For the sake of clarity, we limit our analysis to $N = 3$, but the mechanisms behind under-employment are general and would be present with more islands or worker types.

When workers can respond to the presence of higher-skill individuals, the equilibrium level of under-employment is determined by the interaction of two forces, instead of just one in the $N = 2$ case: (i) a force that "pushes" workers down the occupation ladder: as high-skill workers invade the island below, they push lower-skill individuals further down the ladder, exactly as in the $N = 2$ case, and (ii) a force that "pulls" workers down the ladder: as low-skill workers move down the ladder, they free up space in their island, which pulls the higher-skill individuals down the ladder.

We first characterize the equilibrium allocation and then present some comparative statics exercises to illustrate the mechanisms underlying the equilibrium.

Equilibrium The equilibrium with three types of workers and firms is characterized by the following Proposition:

Proposition 2. *With $N = 3$, there is a unique equilibrium allocation of workers satisfying*

- *Type 3 workers are indifferent between islands 3 and 2, and x_3, the share of type 3 workers searching in island 2, is given by the arbitrage condition*

$$A_3(x_3, x_2) = -E\omega_{3,3} + E\omega_{3,2} = 0 \quad (A_3)$$

with

$$\begin{cases} E\omega_{3,3} = \beta e^{-q_3(1-x_3)} p_3 \phi_3 \\ E\omega_{3,2} = \beta e^{-q_2 x_3 h_3} e^{-q_2(1-x_2)} p_3 \phi_2 + \beta e^{-q_2 x_3 h_3} \left[1 - e^{-q_2(1-x_2)}\right] \phi_2 (p_3 - p_2) \end{cases}$$

- *Type 2 workers are indifferent between islands 2 and 1, and x_2, the share of type 2 workers searching in island 1, is given by the arbitrage condition*

$$A(x_3, x_2) = -E\omega_{2,2} + E\omega_{2,1} = 0 \quad (A_2)$$

with

$$\begin{cases} E\omega_{2,2} = \beta e^{-q_2(1-x_2) - q_2 x_3 h_3} p_2 \phi_2 \\ E\omega_{2,1} = \beta e^{-q_1 x_2 h_2} e^{-q_1} p_2 \phi_1 + \beta e^{-q_1 x_2 h_2} \phi_1 (p_2 - p_1) \left[1 - e^{-q_1}\right] \end{cases}$$

- *Type* 1 *workers only look for jobs in island* 1.

Proof. Appendix. □

As illustrated in figure 9, the two arbitrage equations (A_3) and (A_2) implicitly define a unique equilibrium allocation (x_2, x_3). As shown in the Appendix, both curves are increasing, but the (A_2) curve is always steeper than the (A_3) curve.

To get some intuition, recall that the (A_2) curve captures the decision of type 2 workers to search in island 2 or 1. The (A_2) curve is increasing, because an increase in x_3, the number of type 3 workers in island 2, raises congestion in island 2, which "pushes" type 2 workers down to island 1 and increases x_2. The (A_3) curve captures the decision of type 3 workers to search in island 3 or 2. The (A_3) curve is increasing, because an increase in x_2, the number of type 2 workers in island 1, lowers congestion in island 2, which attracts, i.e., "pulls", type 3 workers down to island 2 and increases x_3. The fact that the (A_2) curve is always steeper than the (A_3) curve means that the "pushing effect" is always stronger than the "pulling effect".

Mechanisms Compared to the $N = 2$ case, under-employment is determined by the interactions of two forces: (i) a force that "pushes" workers down the ladder, captured by the (A_2) curve: as higher type workers invade the island below, they push the lower types further down the ladder, as in the $N = 2$ case, and (ii) a force that "pulls" workers down the ladder, captured by the (A_3) curve: as the lower types move down the ladder, they free up space in their islands, which pulls the higher types even further into their island.

In order to understand how these two forces interact, consider the thought experiment in which island $n = 1$ was closed for agents of type $n = 2, 3$. This initial point corresponds to the point E_0 in figure 9 and is identical to the $N = 2$ case previously discussed: type 2 agents are stuck in island 2 and $x_2 = 0$. Imagine that island 1 suddenly opens up, allowing anyone to look for a job in island 1.

1. Given x_3^0, the initial fraction of type 3 workers in island 2, workers in island 2 have an incentive to look for a job in island 1, because E_0 is above the (A_2) curve, so that $(A_2(0, x_3)) > 0$ and $E\omega_{2,1} > E\omega_{2,2}$. As a result a fraction x_2^1 of type 2 workers moves down to island 1, up until the point where $(A_2(x_2^1, x_3)) = 0$ (point E_1 in figure 9). In effect, type 2 workers are "pushed down" the ladder by type 3 workers, and this "pushing" effect is captured by the curve (A_2).

2. Following the downward movement of type 2 workers, island 2 is less congested than when type 3 agents initially made their island choice, and E_1 is below

the (A_3) curve, so that $E\omega_{3,3} < E\omega_{3,2}$. As a result, more type 3 workers will descend to island 2 up until the point where $(A_3(x_2^1, x_3^2)) = 0$ with x_3^2 the new number of type 3 workers in island 2 (point E_2 in figure 9). In effect, type 3 workers are "pulled down" the ladder by type 2 workers leaving their island, and this "pulling" effect is captured by the curve (A_3).

3. Again, type 2 workers respond to the increased number of type 3 workers by further descending down to island 1, which triggers a response from type 3 workers and so on. This cascade ends at the equilibrium point E.

Comparative Statics: the Effect of Job Polarization We now discuss one comparative statics exercise to illustrate how the interactions between agents' decisions across islands (when $N > 2$) play out in equilibrium, and how a local shock can end up affecting all workers.

Consider an adverse labor demand shock hitting the middle productivity island, i.e., an increase in the queue length q_2. This thought experiment can be seen as studying the effect of job polarization and the disappearance of jobs in middle-skill occupations (the "hollowing out" of the skill distribution, Autor (2010)) on the allocation of workers.

Job polarization has two effects (see figure 10). On the one hand, the (A_2) curve shifts down, because of fewer job opportunities for type 2 workers in island 2, which would increase x_2, i.e., under-employment. On the other hand, the (A_3) curve also shifts down, because of fewer job opportunities for type 3 workers in island 2. This decreases x_2, because there are fewer type 3 workers pushing type 2 workers down the ladder. Overall, the effect of job polarization on the under-employment rate of middle-skill workers is thus ambiguous. However, under-employment amongst high-skill workers will unambiguously decrease.

In terms of expected income, it is easy to show that job polarization leads the expected income of type 3 (high-skill) workers to decrease and the expected income of type 2 (middle-skill) workers to decrease. However, the expected income of type 1 (low-skill) workers can either increase or decrease, depending on the effect of job polarization on the under-employment rate of middle-skill workers.

3 General equilibrium with Endogenous Labor Demand

In this section, we characterize the general equilibrium with endogenous labor demand.

There is an arbitrarily large mass of potential entrants who can settle in island

n by adopting technology ϕ_n. A firm consists of one vacancy, and a firm can enter an island n by posting a vacancy at a cost $c_n > 0$. With free entry, firms will enter in each island n until the point where expected profits, denoted π_n, equal the fixed cost c_n. The number of firms/vacancies in each island will thus be determined endogenously by firm entry.[12]

Definition 2. *General Equilibrium Allocation with Endogenous Firm Entry.*

Each worker of each type $i \in \{1, \ldots, N\}$ decides on which island $n \in \{1, \ldots, N\}$ to search for a job, and each potential firm entrant in island $j \in \{1, \ldots, N\}$ decides whether or not to post a vacancy in island n. The equilibrium is given by a sequence of choices $C_i : [0, L_i] \mapsto \{1, \ldots, N\}$ for each worker i, a sequence of choices $F_j : [0, \infty) \mapsto \{0, 1\}$ for each potential entrants in island j, a matching function between applicants and firms and a sequence of allocations and wages.

3.1 Endogenous firm entry with $N = 2$: two worker types and two firm types

We start with the simplest $N = 2$ framework with 2 types of workers and 2 islands, as studied in Section 3.

Equilibrium The equilibrium is characterized by the following Proposition:

Proposition 3. *With $N = 2$, there is a unique equilibrium allocation satisfying*

- *The arbitrage conditions characterizing the allocation of workers*
  - *Type 2 workers are indifferent between islands 1 and 2, and x_2, the share of type 2 workers searching in island 1, is given by the arbitrage condition*

$$A(x_2, q_1) = -E\omega_{2,2} + E\omega_{2,1}(x_2, q_1) = 0 \qquad (L^S)$$

  - *Type 1 workers only look for jobs in island 1: $x_1 = 0$.*
- *Firms' free entry conditions (market clearing) in islands 1 and 2*

$$\begin{cases} \pi_1(x_2, q_1) = c_1 & (L_1^D) \\ \pi_2(x_2, q_2) = c_2 & (L_2^D) \end{cases}$$

[12] Note the parallel between our model with endogenous free entry and the competitive search model (Moen (1997)), where firms can commit to their posted wage: While competition occurs over prices (i.e., wages) in competitive search models, in our model, competition occurs over quantities: Free entry and the resulting no-profit condition determines the quantity of firms in each island.

Proof. Appendix. □

The expected profit for a firm with technology 1 is given by

$$\pi_1(x_2, q_1) = p_2\phi_1 - \phi_1 e^{-x_2 h_2 q_1}\left[\left(p_2 - p_1 + p_1 e^{-q_1}\right)\left(1 + (2-\beta)x_2 h_2 q_1\right) + (2-\beta)p_1 q_1 e^{-q_1}\right]$$

and the expected profit for a firm with technology 2 is

$$\pi_2(x_2, q_2) = \left(1 - e^{-(1-x_2)q_2} - (1-x_2)q_2 e^{-(1-x_2)q_2}\right)p_2\phi_2 + (1-x_2)q_2 e^{-(1-x_2)q_2}(1-\beta)p_2\phi_2.$$

The general equilibrium allocation is thus the triple (x_2, q_1, q_2) determined by firms' free entry conditions in islands 1 and 2, and the arbitrage equation between islands 1 and 2 for type 2 workers. To gain some intuition about the mechanics of the general equilibrium allocation, it is helpful to depict the allocation graphically. We first represent graphically the worker allocation problem, similarly to Section 2, and study how general equilibrium forces affect the partial equilibrium allocation. We then represent graphically the equilibrium allocation in island 1.

Endogenous labor demand and the expected income of type 2 Similarly to figure 6 in the partial equilibrium case, figure 11 depicts the equilibrium allocation of type 2 workers as the intersection the $E\omega_{2,1}$ curve, the expected wage earned in island 1, and the $E\omega_{2,2}$ curve, the expected wage earned in island 2. To illustrate how general equilibrium forces (endogenous labor demand) affects the equilibrium allocation of workers and the extent of under-employment, figure 11 also plots (dashed lines) the expected wage schedules in partial equilibrium. We can see that competition between firms (i) makes the $E\omega_{2,2}$ curve horizontal and (ii) flattens the $E\omega_{2,1}$ curve.

The following corollary captures formally how the expected wage in island 1 or 2 depends on the share of type 2 workers searching in island 1.

Corollary 1. *The expected income of type 2 workers searching in island 2, $E\omega_{2,2}(x_2, q_2(x_2))$, is independent of x_2. The expected income of type 2 workers searching in island 1, $E\omega_{2,1}(x_2, q_1(x_2))$, is strictly decreasing in x_2 with $\left|\frac{dE\omega_{2,1}}{dx_2}\right| < \left|\frac{\partial E\omega_{2,1}}{\partial x_2}\right|$.*

Proof. Appendix. □

We now discuss the intuition behind this result.

Consider first the $E\omega_{2,2}$ curve, the expected wage earned in island 2. With free entry, the wage schedule becomes flat, and the number of job seekers has no effect on the expected wage. This result comes from the fact workers are homogenous in island 2 (there are only type 2 workers), so that free entry pins down the equilibrium queue length in island 2 –$q_2(1-x_2)$–, regardless of the number of type 2 workers (i.e., regardless of $1-x_2$). This can be seen easily from (L_2^D) which uniquely pins down $q_2(1-x_2)$. This is similar to what happens in standard search and matching models (Mortensen and Pissarides (1994)) where the supply of (homogenous) labor has no effect on the equilibrium queue length (Pissarides (2000)). Even though a higher number of type 2 workers improves the matching probability of a firm, free entry ensures that more firms enter the market in order to keep profits constant.[13]

Consider now the $E\omega_{2,1}$ curve, the expected wage earned in island 1. Compared to the partial equilibrium case with exogenous labor demand, the wage schedule is flatter with endogenous firm entry: an increase in the number of type 2 workers has a smaller negative effect on the expected wage of type 2 workers in island 1. This general equilibrium effect comes from the fact that with free entry, the queue length q_1 is an endogenous variable that responds to the presence of type 2 workers, i.e., q_1 is a function of x_2: $q_1(x_2)$. An inflow of type-2 workers in island 1 raises firms' profits, because firms generate a higher profit when hiring type 2 workers than when hiring type 1 workers. This fosters entry and limits the increase in congestion generated by the inflow of workers. With a muted increase in congestion, workers' bargaining power does not decrease as fast in general equilibrium as in partial equilibrium, and the wage curve is less steep.

Endogenous labor demand and the expected income of type 1 In partial equilibrium, an increase in the number of type 2 workers in island 1 unambiguously lowers the expected income of type 1 workers $E\omega_{1,1}$ (figure 12). This is no longer necessarily the case in general equilibrium. As illustrated in figure 12, with endogenous labor demand, the expected income of type 1 workers may actually increase with the number of type 2 workers. The following corollary captures this result formally:

Corollary 2. *The expected income of type 1 workers searching in island 1, $E\omega_{1,1}(x_2, q_1(x_2))$,*

[13]In a search and matching model, at a given vacancy level, an increase in the number of job seekers (coming from say out of the labor force, as in Pissarides (2000), Chapter 5) raises firms matching probability, i.e., reduces hiring costs, and leads more firms to enter the market, keeping profit and thus the queue length unchanged. A difference between our framework and the Mortensen-Pissarides model is that, in our set-up, an increase in the supply of workers also improves the bargaining position of the firm (as workers compete against each other when negotiated the wage). This difference has no consequence on the equilibrium queue length, because the bargaining position is also solely a function of the queue length $q_2(1-x_2)$. As a result, no matter the level of $1-x_2$, free entry ensures that the queue length adjusts to keep profits (including the fix cost) nil. Importantly, we will see that this is no longer be the case with heterogeneous applicants, and that the bargaining position of the firm is a key determinant of the equilibrium allocation.

is a non-monotonic function of x_2*; decreasing over* $[0, x_2^*]$ *and increasing over* $[x_2^*, 1]$ *with* $x_2^* \in [0, 1]$.

Proof. Appendix. $\square$

The reason for this result is similar to the reason why the $E\omega_{2,1}$ curve is flatter in general equilibrium: an increase in the number of type 2 workers raises firms' profits and leads to more job creation. If the increase in job creation is strong enough, it can more than compensate the increase in congestion generated by the arrival of type 2 workers.

To get some intuition about why $E\omega_{1,1}$ is initially decreasing with x_2 and then increasing, recall that the only instance in which a type 1 worker actually gets a positive surplus is when he is the only applicant. The likelihood of this configuration depends on the number of type 2 workers, and specifically on the probability of *not facing* a type 2, which is given by $P(a_2 = 0) = e^{-q_1 x_2 h_2}$. Because of the convexity of that function of x_2, the adverse effect of an additional type 2 on the expected income of a type 1 is strongest for low values of x_2 and weakest for high values of x_2. In contrast, firms' incentive to post more vacancies depends on the likelihood of *facing* a type 2 applicant, , which is given by $P(a_2 > 0) = 1 - e^{-q_1 x_2 h_2}$, and is thus strongest for high values of x_2. As a result, the general equilibrium effect is strongest for high values of x_2 and $E\omega_{1,1}$ is increasing when x_2 is large enough.

Graphical representation of equilibrium allocation We can now graphically represent the equilibrium allocation. Recall that the general equilibrium allocation is determined by the triple (x_2, q_1, q_2). Since free entry in island 2 fixes $q_2(1 - x_2)$, characterizing the equilibrium allocation reduces to finding the pair (x_2, q_1) that satisfies (i) firms' free entry condition in island 1 and (ii) type 2 worker's arbitrage condition. Although one could depict the equilibrium in the (x_2, q_1) space, we prefer to depict it in the (x_2, V_1) space (recall that $V_1 = L_1/q_1$ with L_1 fixed), since it corresponds to the (U, V) space representation used in standard search and matching models.

As shown in figure 13, the equilibrium is then determined by the intersection of two curves: a "labor demand curve", (L_1^D), given by firms' free entry condition (also called job creation condition) as in search and matching models, and a "labor supply curve", (L^S), characterizing the number of type 2 workers in island 1 and given by the arbitrage condition of type 2 workers between islands 1 and 2.

The labor demand curve is upward sloping and non-linear. To understand the shape of the labor demand curve (L_1^D), it is useful to go back to the standard

Mortensen-Pissarides (MP) model, in which workers are homogeneous. Recall that the total number of job seekers in island 1 is given by $L_1(1 + x_2h_2)$. We can thus represent the labor demand curve, or job creation curve, in a similar fashion to MP models by plotting the job creation curve in (U, V) space. Starting from a world with only type 1 workers and $x_2 = 0$ (i.e., being to the left of the y-axis in figure 13), all workers are homogeneous and, as in the MP model, increasing the number of type 1 (increasing L_1) does not affect the equilibrium queue length V_1/L_1. As a result, the labor demand curve (dashed blue line) crosses the origin at 0. Now, consider the case where one adds type 2 workers and $x_2 > 0$. Because firms generate a higher profit when hiring type 2 workers than when hiring type 1 workers, an increase in x_2 generates a disproportionate increase in the number of firms in island 1, and the equilibrium queue length $\frac{V_1}{L_1(1+x_2h_2)}$ increases. In other words, the slope of the labor demand curve is initially increasing with x_2. This portion of the labor demand curve can be seen as capturing a "quality effect": as the share of type 2 workers in island 1 increases, the quality (i.e., skill level) of the average applicant improves, and this leads to a disproportionate increase in job creation. Then, as the number of type 2 workers becomes large relative to the number of type 1, the labor market in island 1 resembles more to more to that of an homogeneous market with only type 2 workers, in which the queue length is independent of the number of type 2 and the slope of (L_1^D) is again independent of x_2 (dashed red line).

The labor supply curve is capturing how x_2 depends on V_1 and is also upward slopping: the larger the number of job openings, the less competition type 2 workers will face when searching in island 1, and the higher their expected wage. As a result, an increase in V_1 raises the incentive of type 2 to move down to island 1 and increases x_2. The labor supply curve can be seen as capturing a "pulling" effect, similar to the one discussed in Section 3: when the number of job openings in the lower tech island (island 1) increases, it lowers congestion in island 1, i.e., frees up space in island 1, which "pulls" the higher types down to the lower tech island island.

Comparative statics We consider two comparative statics exercises: (i) an adverse labor demand shock affecting the high productivity island, and (ii) an adverse labor demand shock affecting the low productivity island. We will see that underemployment generates an asymmetry in the effects of the two shocks on the labor markets of the high-skills and the low-skills: While high-skill workers are never affected by shocks hitting the low productivity island, low-skill workers are always affected by shocks hitting the high productivity island. Moreover, we will see that, compared with the partial equilibrium case, general equilibrium effects dampen the effect of a shock in the high productivity island on low-skill workers, but exacerbate the effect of a shock in the low productivity island on low-skill workers. In both cases, the general equilibrium effect comes from the fact that changes in the under-employment rate of high-skill workers affect the quality of the average applicant and thus firms' job creation decision.

First, similarly to the comparative statics exercise in partial equilibrium, we now consider an adverse labor demand shock affecting the high productivity island, i.e., an increase in the vacancy posting cost c_2. As shown in the left panel of Figure 14, an increase in c_2 shifts right the labor supply curve, as type 2 workers have a stronger incentive to look for job opportunities in island 1. If the labor demand curve was exogenously fixed (and thus horizontal) as in the partial equilibrium case, the increase in x_2 would be limited to point B. However, because both the labor demand and labor supply curves are upward slopping, the reduction in the number of type 2 workers present in island 1 reduces job openings, which further reduces the number of type 2 workers in island 1, which in turn leads to a stronger decline in job openings, etc.. This general equilibrium effect leads to point C, in which both job openings and under-employment are higher. Interestingly, while in partial equilibrium, a negative labor demand shock in island 2 unambiguously lowers the expected income of type 1 workers, this is no longer the case in general equilibrium. As we saw in in Figure 12, the trickle down of unemployment can actually be good for low-skill workers: a negative labor demand shock in island 2 can actually increase the income of type 1 workers, thanks to the increased job creation triggered by the improvements in the quality of the applicant pool in island 1.

Although the notion that unemployment may trickle-down to the lower layers suggests that under-employment would exacerbate inequality, this comparative statics exercise shows that this is not necessarily the case in general equilibrium. Because an increase in the number of high-skill workers in low-technology islands raises the average quality (i.e., productivity) of the unemployment pool, under-employment stimulates job creation, which indirectly benefits low-skill workers. As a result, a negative labor demand shock in the high-technology sector can actually lower wage inequality between high-skill and low-skill workers.

Finally, note that the partial/general equilibrium distinction can be interpreted as a short-run/long-run distinction. In the short-run, firms cannot respond to the quality and size of the unemployment pool. As a result, immediately after an adverse labor demand shock, the descent of high-skill workers to low productivity islands hurts low-skill workers and increases wage inequality. However, in the long-run, firms react to the presence of higher-skill workers by posting more vacancies, and that may ultimately benefit low-skill workers. Naturally, depending on the time needed to reach this long-run equilibrium, the cost of business fluctuations may still be born disproportionately by low-skill workers.

Second, we consider an adverse labor demand shock in island 1, i.e., an increase in c_1 shifts down the labor demand curve (right panel, Figure 14). If the labor demand curve was exogenously fixed (and thus horizontal) as in the partial equilibrium case, the drop in job openings would be limited to point B. However, because both the labor demand and labor supply curves are upward sloping, the reduction in job openings reduces the number of type 2 workers present in island 1. This in turn leads to a stronger decline in job openings, which further reduces the number of type 2 workers in island 1, etc.. This general equilibrium effect leads to point C, in which both job openings and under-employment are lower than in the partial equilibrium outcome B. In other words, the adverse consequence of a negative labor demand shock in island 1 on type 1 workers is magnified by the deterioration in the average quality of the pool of applicant, and type 1 workers suffer a larger drop in income. Finally, note that despite movements in x_2, the income of type 2 workers was never affected and $E\omega_{2,2}$ remained constant.

3.2 Endogenous firm entry with $N = 3$: three worker types and three firm types

The equilibrium with three types of workers and firms is characterized by the following Proposition:

Proposition 4. *With $N = 3$, there is a unique equilibrium allocation satisfying*

- *The arbitrage conditions characterizing the allocation of workers*

  - *Type 3 workers are indifferent between islands 2 and 3, and x_3, the share of type 3 workers searching in island 2, is given by the arbitrage condition*

$$A(x_3, x_2, q_2) = -E\omega_{3,3} + E\omega_{3,2}(x_3, x_2, q_2) = 0$$

  - *Type 2 workers are indifferent between islands 1 and 2, and x_2, the share of type 2 workers searching in island 1, is given by the arbitrage condition*

$$A(x_3, x_2, q_2, q_1) = -E\omega_{2,2}(x_3, x_2, q_2) + E\omega_{2,1}(x_2, q_1) = 0$$

  - *Type 1 workers only look for jobs in island 1: $x_1 = 0$.*

- *Firms' free entry conditions (market clearing) in islands 1, 2 and 3*

$$\begin{cases} \pi_1(x_2, q_1) = c_1 \\ \pi_2(x_3, x_2, q_2) = c_2 \\ \pi_3(x_3, q_3) = c_3 \end{cases}$$

Proof. Appendix. □

The general equilibrium allocation with $N = 3$ is thus the vector $(x_3, x_2, q_3, q_2, q_1)$ determined by firms' free entry conditions in islands 1, 2 and 3, and the arbitrage equations for type 2 and type 3 workers.

4 Constrained optimal allocation

We now consider the efficiency property of the decentralized allocation. We study the problem of a planner who can only allocate workers across islands in order to maximize total output (net of the cost of posting vacancies), while taking into account firms' free entry condition. Importantly, the planner is constrained by the existence of coordination frictions in each island and by the impossibility for firms to commit to a posted wage.

We first consider the $N = 2$ case, and then study the effect of interactions across islands in the more general $N = 3$ case.

4.1 Efficiency with $N = 2$: two worker types and two firm types

The following proposition states the decentralized allocation is, in general, not efficient, and that there is too much under-employment.

Proposition 5. *When $N = 2$, the constrained optimal allocation (x_2^*, q_1^*) is characterized by the firms' free entry conditions in islands 1 and 2, and*

$$\begin{aligned} A(x_2^*, q_1^*) &= -E\omega_{2,2} + E\omega_{2,1} \\ &= (1-\beta)\, h_2 q_1 \phi_1^2 p_1 e^{-q_1^*(2x_2^* h_2 + 1)}\, (p_2 - p_1)\, \frac{1}{\frac{\partial \pi_1(x_2^*, q_1^*)}{\partial q_1}} \qquad (1) \\ &\geq 0 \end{aligned}$$

with the expression for $\frac{\partial \pi_1(x_2, q_1)}{\partial q_1} > 0$ given in the appendix.

If $\beta < 1$ and $p_2 - p_1 > 0$, the decentralized allocation (x_2, q_1) is inefficient and has too much under-employment: $x_2 > x_2^*$.

Proof. See appendix. □

Contrasting the decentralized and centralized allocations To better understand the externalities present in this economy, it is useful to contrast the worker's problem and the planner's problem. Type 2 workers allocate themselves between islands 1 and 2 up until

$$E\omega_{2,2} = E\omega_{2,1}. \tag{2}$$

In contrast, the planner wishes to allocate workers in order to maximize total output, while satisfying firms' zero profit condition. With free entry, we have $\pi = y - \omega = c$ so that maximizing total output is identical to maximizing the total wage bill Ω. The planner's problem is thus to maximize the wage bill while satisfying firms' free entry conditions

$$\begin{cases} \max_{x_2} \Omega \\ \Omega = (1-x_2)h_2 E\omega_{2,2} + h_2 x_2 E\omega_{2,1}(x_2, q_1(x_2)) + E\omega_{1,1}(x_2, q_1(x_2)) \end{cases} \tag{3}$$

with $q_1(x_2)$ is given by fims' free entry condition in island 1: $\pi_1(x_2, q_1(x_2)) = c_1$.

Finally, note that we omitted one externality in our discussion: since type 2 workers do not internalize how their descent affects $E\omega_{2,2}$, the wage of high-skill workers who remained in the high-skill island, the marginal high-skill applicant should also exert an externality on $E\omega_{2,2}$. This externality has no consequence however, because $E\omega_{2,2}$ is constant thanks to general equilibrium effects. Indeed, and implicit in the planner's problem(3), we used the fact that the free entry condition in island 2 ensures that $E\omega_{2,2}$ is constant and independent of x_2 and q_1.

Contrasting (2) with (3), we can see that type 2 individuals are solving the planner's problem treating $E\omega_{2,1}(x_2, q_1(x_2))$ and $E\omega_{1,1}(x_2, q_1(x_2))$ as independent of x_2. In other words, type 2 workers do not internalize how their descent affect the wages of other workers , i.e., how an increase in x_2, the fraction of type 2 workers in island 1, affects (i) competition between types, (ii) competition across types, and (iii) job creation in island 1.

1. *directly* affects the wage of type 2 workers $E\omega_{2,1}(x_2, .)$: a higher share of type 2 workers raises competition between type 2 workers, which lowers the expected wage $E\omega_{2,1}$. This is a (negative) within-type externality.

2. *directly* affects the wage of type 1 workers $E\omega_{1,1}(x_2, .)$: a higher share of type 2 workers raises the quality of the competition faced by type 1 workers, which lowers $E\omega_{1,1}$, the expected wage of type 1 workers. This is a (negative) across-type externality.

3. *indirectly* affects wages through firms' job creation condition: $E\omega_{2,1}(., q_1(x_2))$

and $E\omega_{1,1}(., q_1(x_2))$: a higher share of type 2 workers raises firm's profit, which leads through the no-profit condition to more job openings, and thus to higher wages for both type 1 and type 2. We label this (positive) externality a job creation externality.

Intuition: the ranking externality

To get some intuition about the (in)efficiency property of the decentralized allocation, the efficiency discussion can be summarized in terms of one single externality, that we label a *ranking externality*.

The key element behind the ranking externality is the fact that high-skill workers and low-skill workers face *different* labor markets in the *same* island: with non-random hiring, a high-skill worker always gets the job when in competition with a low-skill worker, because he is ranked first. As a result, a high-skill worker's decision to move down the occupation ladder are motivated by different factors than those determining the labor market prospect of low-skill workers. Consider for instance a labor market with many low-skill workers but no high-skill worker. In that case, high-skill and low-skill face completely opposite labor market prospects: the market is very congested (and unattractive) for low-skill workers, but it is very attractive for high-skill workers, because they face no competition.

Since the gain from moving down the ladder for a high-skill is strongly disconnected from the loss inflicted on the other workers, the decision to move down the ladder is disconnected from the loss inflicted on the other workers, and the ranking externality can be large. The magnitude of the ranking externality depends on how a marginal high-skill applicant differs from the average applicant in the low-productivity island. The externality is strongest (as in our example) when the marginal high-skill applicant is most different from the average applicant, i.e., most "unique".

To see this more formally, we can contrast (2) with (3). The decentralized allocation is efficient if and only if

$$\frac{d\Omega_1}{dx_2} \equiv \frac{dE\omega_{1,1}(x_2, q_1(x_2))}{dx_2} + x_2 h_2 \frac{dE\omega_{2,1}(x_2, q_1(x_2))}{dx_2} = 0. \tag{4}$$

Expression $\frac{d\Omega_1}{dx_2}$ captures the ranking externality, i.e., the effect of a marginal high-skill worker searching in the low-tech island on the labor markets of (i) low-skill workers ($\frac{dE\omega_{1,1}}{dx_2}$) and (ii) high-skill workers ($x_2 h_2 \frac{dE\omega_{2,1}}{dx_2}$) weighted by their relative population shares. In essence, it captures the congestion externalities net of the compensating effect of the job creation externality.

Using the expression for $\frac{d\Omega_1}{dx_2}$, one can visualize in figure 15 how the externality evolves with the presence of high-skill workers in the low-tech island. Starting from a world with no high-skill workers in the low-tech island ($x_2 = 0$), the presence of one high-skill worker imposes a large cost to low-skill workers: since a high-skill is systematically ranked above competing low-skill applicants, the high-skill applicant systematically gets the job when in competition with a lower-skill applicant, and the labor market of the low skills deteriorates sharply ($\frac{dE\omega_{1,1}}{dx_2} < 0$).

As the number of high-skill workers increases, high-skill workers become less unique and the high-skills start competing against each other for their premium over the low-skills. As each high-skill worker becomes less unique, his wage premium over an average applicant deteriorates, and the ranking externality, which stems from the difference between a marginal high-skill and the average worker in island 1, becomes less strong.[14]

As x_2 increases further, the job creation externality becomes stronger (because firms are more likely to face a high-skill applicant) and compensates the increased congestion in the labor market. In fact, the income of low-skill workers starts increasing for x_2 large enough. However, this cannot compensate the increased congestion between the high-skill workers, and the ranking externality remains negative.[15] Intuitively, job creation can never fully compensate the congestion externality, because, while firms' profits increase with the number of high-skill workers, the presence of low-skill workers limits the increase in profits, because firms always face a non-zero probability of receiving just one application from a low-skill applicant and then ending up with a low-skill worker and low profit.

In fact, it is only when the number of high-skills becomes arbitrarily large compared to the number of low-skills, that job creation exactly compensates the increased congestion. Specifically, as $x_2 h_2$ increases further and becomes arbitrary large (if h_2 is arbitrary large), the labor market in the low-tech island resembles that of an homogeneous labor market with only high skill workers, and as in search models with homogeneous labor, the marginal high-skill applicant has no effect on

[14]Note also that as x_2 increases, a marginal high-skill applicant hurts relatively less the low-skills. Indeed, the only case making $E\omega_{1,1}$ positive is the case where the low-skill worker is the only applicant. The likelihood of this configuration (from the perspective of a type 1 worker) is $P(a_2 = 0) = e^{-q_1 x_2 h_2}$, which decreases less and less fast as x_2 increases. As a result, a marginal type 2 has a smaller and smaller effect on the labor market of type 1 workers.

[15]Intuitively, the job creation externality works for both types, but it makes only $E\omega_{1,1}$ increasing (and not $E\omega_{2,1}$) because $E\omega_{1,1}$ is already at a very low level compared to $E\omega_{2,1}$. While the only case making $E\omega_{1,1}$ positive is the case where the low-skill worker is the only applicant to a job, a high-skill worker earns a wage premium, (in part) because he systematically gets the job when in competition with a lower-skill. This case becomes less likely as x_2 increases, which lowers the $E\omega_{2,1}$ and prevents job creation from making $E\omega_{2,1}$ increasing.

the equilibrium queue length (as we describe in more detail below) and $\frac{d\Omega_1}{dx_2} \to 0$, i.e., the ranking externality converges to zero. Intuitively, with x_2h_2 large, the marginal high-skill applicant is identical to the average unemployed in the low-productivity island, and the ranking externality is nil.

This intuition for why $\frac{d\Omega_1}{dx_2} \to 0$ when island 1 is mainly populated by (homogeneous) high-skill workers also explains why type 2 workers exert no externality on workers in island 2. In that (homogeneous) island, the marginal high-skill applicant is always identical to the average unemployed, and the ranking externality is always nil.

4.2 Conditions for Efficiency of the Decentralized allocation

The decentralized allocation can be efficient when (4) is verified. We now discuss two cases that ensure efficiency by getting rid of the ranking externality.

Homogeneous workers A first and trivial case which satisfies (4) is when workers are homogeneous ($p_2 = p_1$). According to Proposition 4, it implies $A(x_2, q_1) = 0$, and the centralized and decentralized allocations coincide.

When workers are homogeneous, as in the standard Mortensen-Pissarides model, the number of job seekers has no effect on the ratio of job openings to job seekers: the firm responds to the addition of one more worker by creating one more vacancy.[16] As a result, positive and negative externalities exactly cancel out for each type: the congestion generated by the addition of one more worker is exactly compensated by the posting of more vacancy such as to keep the equilibrium queue length unchanged. Mathematically, we have

$$\begin{cases} \frac{dE\omega_{2,1}(x_2,q_1(x_2))}{dx_2} = 0 \\ \frac{dE\omega_{1,1}(x_2,q_1(x_2))}{dx_2} = 0 \end{cases}$$

so that (4) is satisfied and the decentralized allocation is efficient.

Bargaining parameter: $\beta = 1$ According to Proposition 4, giving all the bargaining power to the worker when a worker and a firm negotiate how to share the surplus, i.e., a bargaining parameter $\beta = 1$, ensures $A(x_2, q_1) = 0$, and the centralized and decentralized allocations coincide.

Intuitively, the equilibrium is efficient with $\beta = 1$ because a high-skill worker always receives his expected marginal product, no matter whether he is searching

[16] Mathematically, it is easy to see that, when $p_1 = p_2$, the free entry condition ($\pi_1 = c_1$) pins down $q_1(1+x_2h_2)$, the queue length (the job seekers to job openings ratio) in island 1. As a result, a higher share of type 2 workers is exactly compensated by more job openings.

in the high-tech or in the low-tech island. Consider a high-skill worker applying for a job. When he is not the most productive applicant (i.e., in competition with another high-skill), his marginal product is zero, as he brings nothing to the match. When he is the most productive applicant, he has a positive marginal product, equal to the gap between his productivity and the productivity of the second best applicant, or to the gap between his productivity and the firm's option of remaining vacant, whichever is smaller. With $\beta = 1$, this is exactly the wage he would get. Thus, despite the incomplete markets created by search frictions, the decentralized allocation is constrained efficient.

4.3 Efficiency with $N = 3$: three worker types and three firm types

The following proposition states that the decentralized allocation is generally also inefficient when $N = 3$. In the constrained allocation, there is less under-employment of type 2 workers (lower x_2) and less under-employment of type 3 workers (lower x_2) than in the decentralized allocation.

Proposition 6. *When $N = 3$, the constrained optimal allocation $(x_2, x_3, q_1, q_2, q_3)$ does not coincide with the decentralized allocation. It is characterized by the same free entry conditions in islands 1, 2 and 3 but the difference in expected income between two islands for type 3 and type 2 workers is now respectively*

$$\begin{aligned} A_3(x_2, x_3, q_1, q_2, q_3) &= -E\omega_{3,3} + E\omega_{3,2} \\ &= \frac{(1-\beta)\, h_3 h_2 (1-x_2)^2 q_2 \phi_2^2 p_2 e^{-2q_2 x_3 h_3 - q_2(1-x_2)}\, (p_3 - p_2)}{\frac{\partial \pi_2(x_3, x_2, q_2)}{\partial q_2}} \\ &\geq 0 \end{aligned} \tag{5}$$

and

$$\begin{aligned} A_2(x_2, x_3, q_1, q_2) &= -E\omega_{2,2} + E\omega_{2,1} \\ &= \frac{(1-\beta)\, h_2 q_1 \phi_1^2 p_1 e^{-2q_1 x_2 h_2 - q_1}\, (p_2 - p_1)}{\frac{\partial \pi_1(x_2, q_1)}{\partial q_1}} \\ &+ \frac{(1-\beta)(1-x_2)(p_3 - p_2)\phi_2^2 h_2 p_2 q_2 x_3 h_3 e^{-2q_2 x_3 h_3 - q_2(1-x_2)}}{\frac{\partial \pi_2(x_3, x_2, q_2)}{\partial q_2}} \\ &\geq 0 \end{aligned} \tag{6}$$

with the expression for $\frac{\partial \pi_2(x_3,x_2,q_2)}{\partial q_2} > 0$ and $\frac{\partial \pi_1(x_2,q_1)}{\partial q_1} > 0$ given in the appendix.

Proof. See appendix. ∎

Compared to the $N = 2$ case, interactions across agents' decision introduces an additional effect. Comparing with (1) in the N=2 case, we can notice an additional (positive) term in A_2 in the $N = 3$ case, which brings the constrained allocation further away from the decentralized one. This additional term captures the fact that, when deciding to search in island 1, type 2 workers affect not only the ratio of type 1 to type 2 workers in island 1, which affects the job creation decision of firms in island 1 (as is the $N = 2$ case), but also the ratio of type 2 to type 3 workers in island 2, which affects the job creation decision of firms in island 2. Overall, too many type 2 workers search in island 1, and this leads to too much job creation in both island 1 and island 2.

5 Quantifying the inefficiency

Our model of under-employment makes two key predictions: (i) the decentralized allocation is inefficient and there is too much under-employment, and (ii) under-employment generates an asymmetry in the way high-skill and low-skill workers are affected by labor demand shocks, with high-skill workers being insulated from shocks to low-tech firms, but low-skill workers suffering from shocks to high-tech firms (through the increased number of high-skill workers searching in the low-tech island).

In this section, we seek to test these predictions as well as quantify the magnitude of the inefficiency generated by under-employment. To do so, we rely on our benchmark general equilibrium model with 2 islands.

We first discuss in greater details these two key predictions of our model and propose a simple empirical strategy to test them. We then describe the data, and finally present the empirical results. We find that (i) there is strong support for asymmetric responses to shocks along the occupational ladder, and (ii) the inefficiency is large: a back-of-the-envelope calculation shows that reducing the extent of under-employment by 15 percent would increase total output by about 5 percent.

5.1 Empirical predictions of the model

In our model, under-employment has a non-symmetric effect across worker-skill groups, and the model makes two specific predictions (i) under-employment generates downward pressures on wages in low-tech firms, and (ii) under-employment does not affect wages in high-tech firms.

More specifically, as shown in Proposition 5, when a type 2 (high-skill) worker decides to search in island 1 (the low-tech island):

1. Workers already in island 2 (the high-tech island) are unaffected and their total income (Ω_2) is unchanged:

$$\frac{d\Omega_2}{dx_2} = (1 - x_2)h_2 \frac{dE\omega_{2,2}(x_2, q_2(x_2))}{dx_2} = 0. \tag{7}$$

2. Workers already in island 1 are affected, and their total income (Ω_1) declines:

$$\frac{d\Omega_1}{dx_2} = x_2 h_2 \frac{dE\omega_{2,1}(x_2, q_1(x_2))}{dx_2} + \frac{dE\omega_{1,1}(x_2, q_1(x_2))}{dx_2} \leqslant 0. \tag{8}$$

Importantly, because $\frac{d\Omega_1}{dx_2}$ is not internalized by high-skill workers when deciding to search in the high-tech or low-tech island, $\frac{d\Omega_1}{dx_2}$ captures the strength of the ranking externality.[17] Thus, measuring $\frac{d\Omega_1}{dx_2}$ will allow us to quantifying the magnitude of the inefficiency generated by under-employment.

To test the two predictions (7) and (8), we need to measure $\frac{dE\omega_{2,2}}{dx_2}$, $\frac{dE\omega_{2,1}}{dx_2}$ and $\frac{dE\omega_{2,1}}{dx_2}$, i.e., we need to measure how a marginal type 2 applicant modifies the expected wages of the other workers.

5.2 Empirical strategy

We treat c_1 and c_2 as random variables capturing the labor demand shocks affecting respectively island 1 and island 2. When we test for equation 7 (resp. 8), our aim is to identify variations in x_2 that are exogenous to high-tech wages (resp. low-tech wages). Our theoretical framework provides such variations: A shock on firm entry in the low-tech island (resp. high-tech island), i.e. a change in c_1 (resp. c_2), modifies high-tech wages (resp. low-tech wages) *only* through the induced change in under-employment.

However, we cannot observe directly c_1 or c_2. Instead we observe q_1 and q_2, the "initial" queue lengths in each island (as defined in Section 2), which depend on job openings in each island. To estimate equation 7 using variations in x_2 orthogonal to high-tech wages, we use variations in q_1 in order to predict x_2. In our model, a shock on vacancy costs c_1 moves both under-employment x_2 and job openings in the low-tech island q_1. As such, the variations in x_2 implied by q_1 once we control for q_2 are, in fact, induced by variations in the unobserved underlying cost c_1 and such variations are orthogonal to high-tech wages. To estimate equation 8 using variations in x_2 orthogonal to low-tech wages, we proceed in a similar fashion, but

[17] Note that $\frac{d\Omega_2}{dx_2}$ is also not internalized, but since $\frac{d\Omega_2}{dx_2}$ is zero, this does not generate any inefficiency.

controlling instead for movements in q_1 in order to isolate variations in x_2 driven by variations in c_2.

Mathematically, we estimate the effect of a marginal type 2 worker searching in island 1 on the wages of the other workers in two stages. As a first stage, we estimate the relation

$$x_2 = \beta_1 q_1 + \beta_2 q_2 + \varepsilon \tag{s1}$$

and as a second-stage we test the two predictions (7) and (8) by estimating

$$\begin{cases} \omega_{2,1}\bar{h}_2\bar{x}_2 + \omega_{1,1} = \rho_1\hat{x}_2 + \gamma_1 q_1 + \nu \\ \omega_{2,2}(1-\bar{x}_2) = \rho_2\hat{x}_2 + \delta_2 q_2 + \mu \end{cases} \tag{s2}$$

with $\hat{x}_2$ the fitted-value from the first stage (s1), $\bar{x}_2$ the average share of type 2 searching in island 1 over the sample period, $\bar{h}_2$ the average ratio of type 2 job seekers to type 1 job seekers over the sample period and $\omega_{2,1}$, $\omega_{1,1}$ and $\omega_{2,2}$ the observed wages in islands 1 and 2.[18]

The first equation in (s2) allows us to estimate ρ_1, the effect of changes in x_2 driven by shocks to c_2 on the wages of job seekers in island 1 –a test of prediction (8)–, and the second equation allows us to estimate ρ_2, the effect of changes in x_2 driven by shocks to c_1 on the wages of job seekers in island 2–a test of prediction (7)–.

5.3 Data construction

We use matched CPS micro data between 1994 and 2008 to observe transitions from unemployment into employment. We compute for each observation the worker educational attainment, (1) less than high school or high school –a type 1 worker–, and (2) higher than school, –a type 2 worker–. We associate to each observation the weekly wages that are declared. Finally, we use the 3-digit occupational code that is documented together with the wage and we define three groups of new hires:

- type 1 workers in occupations requiring less than high school (wage $\omega_{1,1}$),
- type 2 workers in occupations requiring less than high school ($\omega_{2,1}$),
- type 2 workers in occupations requiring more than high school ($\omega_{2,2}$).

[18] We can substitute expected wages with observed wages, because observed wages can be written as functions of expected wages: $\omega_{2,2} = f(E\omega_{2,2}, x_2, q_2(x_2))$, $\omega_{2,1} = g(E\omega_{2,1}, x_2, q_1(x_2))$ and $\omega_{1,1} = h(E\omega_{1,1}, x_2, q_1(x_2))$.

Recall that x_2 is the share of type 2 workers applying in island 1. We do not observe directly this quantity, we consider instead the share of type 2 workers *having accepted* a job in island 1.

Finally, we construct measures of vacancy posting in each island (and thus q_1 and q_2), from job openings data by occupation groups constructed by Barnichon and Figura (2013). While vacancies are defined by ocupation groups (professional p, services se, construction c and sales sa) and not by degree requirements, we consider services and professional as occupational categories requiring more than high school (it is true for more than 90% of the 3-digit occupations which compose those two categories), and we consider construction and sales as occupational categories requiring less than high school (it is true for more than 75% of the 3-digit occupations which compose those two categories). Let q_1^p, q_1^{se}, q_2^c and q_2^{sa} denote the ratios $\frac{L}{V}$ as defined in Section 2 for each categories, with L the number of "potential" job seekers in each category, i.e., the number of unemployed who reported that occupation as a previous occupation.

5.4 Results

We first estimate the effect of changes in x_2 driven by shocks to c_2 on the wages of job seekers in island 1 –a test of prediction (8)–. The first two columns of table 1 display the results of the first and second stage. In specifications 1 and 2, both q_2^p and q_2^{se} are used as instruments and proxies for the queue in island 2. In specification 2, we correct for seasonal variations by including monthly dummies. In all cases, the instruments predict well movements in under-employment: the fewer job openings in island 2 (i.e., the larger q_2), the larger the under-employment rate (x_2), but the fewer job openings in island 1, the smaller the under-employment rate. The results of the second stage indicate a weekly wage loss around \$60 per week when x_2 increases by 10 percentage points.

This wage loss is statistically significant, and is consistent with the model prediction (8) that high-skill workers negatively affect the income of job seekers in the low-tech island. Moreover, the estimate gives an indication of the magnitude of the inefficiency. Over the period 1994-2008, the average share of type 2 looking in island 1, $\bar{x}_2$, was about 55% so that the average total weekly wage (i.e., output net of vacancy posting costs) is

$$\underbrace{\Omega}_{\approx 1160\$} = \underbrace{(1-x_2)\omega_{2,2}}_{\approx 400\$} + \underbrace{x_2\omega_{2,1}}_{\approx 360\$} + \underbrace{\omega_{1,1}}_{\approx 400\$}.$$

As a result, decreasing x_2 by 10 percentage points, from 55% to 45%, would increase weekly wages by 60\$, i.e., raise GDP by $\frac{60}{1160} \approx 5\%$, a substantial effect.

Second, we estimate the effect of changes in x_2 driven by shocks to c_1 on the wages of job seekers in island 2 –a test of prediction (7)–. Consistent with the model, the last two columns of table 1 show that, the wage of job seekers in island 2 is unaffected by shocks in island 1 (the low-technology island).

Taking stock, this illustrative empirical exercise shows that the data are consistent with the two key predictions of our model: (i) under-employment generates downward pressures on wages in low-tech firms, and (ii) under-employment does not affect wages in high-tech firms.

6 Conclusion

This paper proposes a theory to study how an individual's decision to move down the occupational ladder can affect the labor market of other workers and thereby affect the functioning of the whole labor market. We present a model of "trickle-down unemployment", in which some high-skill workers try to escape competition from their high-skill peers by moving down the occupational ladder. In doing so, they take the jobs of less-skill individuals, who are in turn driven out their market and down the occupational ladder.

We show that under-employment and the trickle-down of unemployment are generally inefficient. In the decentralized allocation, too many high-skill workers are under-employed, and there are too many low productivity firms and not enough high productivity firms. The existence of a (negative) ranking externality stems from heterogeneity across workers and the fact that higher-skill workers searching for a low productivity job do not internalize how they disproportionately hurt lower-skill workers, because high-skill applicants are systematically ranked first by firms. As a result, the ratio of high-skill to low-skill workers is inefficiently high in the low-productivity island. We show that the larger the difference (or more generally, the variance) between workers' productivity, the stronger the ranking externality.

Importantly, while we emphasized the importance of the ranking externality in the context of under-employment, our results equally apply to any worker allocation problem in which heterogeneous workers compete for the same job. For instance, in the case of "over-employment", in which low-skill workers look for jobs in high-tech islands, it would be easy to show –proceeding as we did in this paper– that the ratio of high-skill to low-skill workers (this time, in the high-skill island) is inefficiently high in the decentralized allocation.

Finally, the possibility of under-employment opens new, and so far unexplored, benefits of education. First, if, ceteris paribus, firms always prefer more educated workers (as is the case in our model), more educated workers have more bargaining power, and they can extract a larger share of a match surplus than less educated workers, i.e., they receive a higher labor income share. Second, a higher education level may not only guarantee a higher expected income, but it may also provide a lower volatility of income, because highly educated workers can partially smooth out adverse labor demand shocks by moving down the occupational ladder. Exploring these additional returns to education is an important goal for future research.

References

A. Gautier, P., J. van den Berg, G., C. van Ours, J. and Ridder, G.: 2002, Worker turnover at the firm level and crowding out of lower educated workers, *European Economic Review* **46**(3), 523–538.

Albrecht, J. and Vroman, S.: 2002, A matching model with endogenous skill requirements, *International Economic Review* **43**(1), 283–305.

Autor, D.: 2010, The polarization of job opportunities in the u.s. labor market., *The Hamilton Project and The Center for American Progress* pp. 1 – 40.

Barnichon, R. and Figura, A.: 2013, Labor market heterogeneity and the aggregate matching function.

Beaudry, P., Green, D. A. and Sand, B. M.: 2013, The great reversal in the demand for skill and cognitive tasks, *NBER Working Papers 18901*, National Bureau of Economic Research, Inc.

Binmore, K., Rubinstein, A. and Wolinsky, A.: 1986, The nash bargaining solution in economic modelling, *The RAND Journal of Economics* **17**(2), pp. 176–188.

Blanchard, O. J. and Diamond, P. A.: 1994, Ranking, unemployment duration, and wages, *Review of Economic Studies* **61**(3), 417–34.

Butters, G.: 1977, Equilibrium distributions of sales and advertising prices, *Review of Economic Studies* **44**, 465–491.

Chang, B.: 2012, A search theory of sectoral reallocation.

Coles, M. G.: 2001, Equilibrium wage dispersion, firm size and growth, *Review of Economic Dynamics* **4**(1), 159–187.

Corominas-Bosch, M.: 2004, Bargaining in a network of buyers and sellers, *Journal of Economic Theory* **115**(1), 35–77.

Doyle, M. and Wong, J.: 2013, Wage posting without full commitment, *Review of Economic Dynamics* **16**(2), 231 – 252.

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

Freeman, R.: 1976, *The Overeducated American*, New York: Academic Press.

Gautier, P. A.: 2002, Unemployment and search externalities in a model with heterogeneous jobs and workers, *Economica* **69**(273), 21–40.

Julien, B., Kennes, J. and King, I.: 2000, Bidding for labor, *Review of Economic Dynamics* **3**(4), 619–649.

Leuven, E. and Oosterbeek, H.: 2011, Overeducation and mismatch in the labor market, *IZA Discussion Papers 5523*, Institute for the Study of Labor (IZA).

Masters, A.: 2005, Directed search without wage commitment and the role of labor market institutions, *Technical report*.

Moen, E. R.: 1997, Competitive search equilibrium, *Journal of Political Economy* **105**(2), 385–411.

Mortensen, D. T. and Pissarides, C. A.: 1994, Job creation and job destruction in the theory of unemployment, *Review of Economic Studies* **61**(3), 397–415.

Pilossoph, L.: 2012, A multisector equilibrium search model of labor reallocation, *Working paper*.

Pissarides, C. A.: 2000, *Equilibrium Unemployment Theory, 2nd Edition*, Vol. 1 of *MIT Press Books*, The MIT Press.

Rubinstein, A.: 1982, Perfect equilibrium in a bargaining model, *Econometrica* **50**(1), 97–109.

Sahin, A., Song, J., Topa, G. and Violante, G. L.: 2012, Mismatch unemployment, *Technical report*.

Shi, S.: 2001, Frictional assignment. i. efficiency, *Journal of Economic Theory* **98**(2), 232–260.

Shi, S.: 2002, A directed search model of inequality with heterogeneous skills and skill-based technology, *Review of Economic Studies* **69**(2), 467–91.

Shimer, R.: 1999, Job auctions.

Shimer, R.: 2005, The assignment of workers to jobs in an economy with coordination frictions, *Journal of Political Economy* **113**(5), 996–1025.

Sicherman, N.: 1991, Overeducation in the labor market, *Journal of Labor Economics* **9**(2), 101–22.

Sicherman, N. and Galor, O.: 1990, A theory of career mobility, *Journal of Political Economy* **98**(1), 169–92.

Spence, A. M.: 1973, Job market signaling, *The Quarterly Journal of Economics* **87**(3), 355–74.

Thurow, L.: 1975, *Generating inequality*, Basic Books.

Table 1. Quantifying the inefficiency.

FIRST STAGE s1	Under-employment x_2			
	(1)	(2)	(1)	(2)
Queue q_2^p	1.68**	1.69**	1.68**	1.69**
	(.279)	(.285)	(.279)	(.285)
Queue q_2^{se}	-.022	-.021	-.022	-.021
	(.023)	(.023)	(.023)	(.023)
Queue q_1^c	-.202**	-.203**	-.202**	-.203**
	(.090)	(.094)	(.090)	(.094)
Queue q_1^{sa}	-.128**	-.130**	-.128**	-.130**
	(.021)	(.022)	(.021)	(.022)
Dummies	-	Month	-	Month
SECOND STAGE s2	$\bar{x}_2\bar{h}_2\omega_{2,1} + \omega_{1,1}$		$(1 - \bar{x}_2)\omega_{2,2}$	
Under-employment x_2	-594.84**	-598.44**	-95.58	-105.62
	(186.94)	(185.04)	(225.05)	(227.26)
Controls (q_1^c, q_1^{sa})	Yes	Yes	-	-
Controls (q_2^p, q_2^{se})	-	-	Yes	Yes
Dummies	-	Month	-	Month
Observations	182	182	182	182
F-stat	21.874	21.205	35.149	33.789

Significantly different than zero at † 90% confidence, * 95% confidence, ** 99% confidence. Standard errors between parentheses are clustered at the year level. Queues are divided by 100 such that $E[q_2^p] = .027$, $E[q_2^{se}] = .152$, $E[q_1^c] = .267$, $E[q_1^{sa}] = .227$. In all specifications, we control for a linear trend.

A Appendix

A.1 Conditions to ensure under-employment in equilibrium

We derive here the conditions that ensure that the equilibrium we consider is an under-employment equilibrium. Our conditions boil down to ensuring that the equilibrium is not at a corner solution in which either everyone or noone is under-employed.

Condition N=2: In the $N = 2$ case, we ensure that some type 2 workers descend to island 1.

In partial equilibrium, this is ensured by:

$$e^{-q_2} p_2 \phi_2 < \left(p_2 - p_1 + p_1 e^{-q_1}\right) \phi_1 \quad (C_2^p)$$

Intuitively, this condition ensures that, when $x_2 = 0$ (no under-employment), the expected wage in island 1 is higher than the expected wage in island 2.

In general equilibrium, we impose a similar condition. We set c_2 such that, in equilibrium, $E\omega_{2,2}(x_2)$ is lower than the expected wage $E\omega_{2,2}(x_2 = 0)$ when there is no under-employment. Formally, the condition can be written as follows:

$$E\omega_{2,2} < \left(p_2 - p_1\right) \phi_1 + E\omega_{1,1}(x_2 = 0) \quad (C_2^g)$$

Condition N=3:

In the $N = 3$ case, we impose conditions guaranteeing (i) there is some under-employment of types 3 and 2, (ii) not all type 2 workers search in island 1 and (iii) type 3 workers do not search in island 1. These conditions ensure that at most 2 types co-exist in a given island.

First, a positive fraction of type 3 and type 2 workers are under-employed as long as:

$$\begin{cases} e^{-q_3} p_3 \phi_3 < \left(p_3 - p_2 + p_2 e^{-q_2}\right) \phi_2 \quad (C_3^p) \\ e^{-q_2} p_2 \phi_2 < \left(p_2 - p_1 + p_1 e^{-q_1}\right) \phi_1 \quad (C_2^p) \end{cases}$$

Second, a positive fraction of type 2 workers search in island 2 as long as:

$$e^{-q_2 h_3} p_2 \phi_2 > e^{-q_1 h_2} \left(p_2 - p_1 + p_1 e^{-q_1}\right) \phi_1 \quad (D_2^p)$$

This condition implies that, even with all type 3 workers in island 2, type 2 workers would not all descend to island 1.

Finally, we derive the condition under which type 3 workers have no incentives to search in island 1. Consider the equilibrium allocation verifying $A_3(x_3, x_2) = 0$ and $A_2(x_2, x_1) = 0$. The expected wage of a type 3 worker searching in island 1 would be

$$E\omega_{3,1} = \phi_1 (p_3 - p_2) + \phi_1 e^{-q_1 x_2 h_2} \left[p_2 - p_1 + p_1 e^{-q_1}\right]$$

Since $\phi_1 e^{-q_1 x_2 h_2} \left[p_2 - p_1 + p_1 e^{-q_1}\right] = e^{-q_2(1-x_2)-q_2 x_3 h_3} p_2 \phi_2$,

$$E\omega_{3,1} = \phi_1 (p_3 - p_2) + e^{-q_2(1-x_2)-q_2 x_3 h_3} p_2 \phi_2$$

In contrast, the expected wage of a type 3 worker searching in island 2 is:

$$\phi_2 e^{-q_2 x_3 h_3} \left[p_3 - p_2 + p_2 e^{-q_2(1-x_2)}\right]$$

It is then immediate that no type 3 workers have the incentives to descend to island 1 as long as:

$$e^{-h_3 q_2} > \tfrac{\phi_1}{\phi_2} \quad (D_3^p).$$

A.2 Proofs

Proof. Proposition 1

Consider first the problem of type 2 workers. A type 2 worker has two choices, he can (i) look for a job in island 2, his "home island", or (ii) look for a job in island 1, i.e., move down the occupation ladder. We now consider these two possibilities.

When a type 2 worker looks for a job in island 2, he faces two possible outcomes: (a) with probability $e^{-q_2(1-x_2)}$, he is the only applicant and receives $\beta p_2 \phi_2$, or (b), with probability $1 - e^{-q_2(1-x_2)}$, he is in competition with other workers and receives 0 (regardless of whether he ends up employed or unemployed). The expected payoff of a worker type 2 who searches for a job in island 2, $\omega_{2,2}$, is thus

$$E\omega_{2,2} = \beta e^{-q_2(1-x_2)} p_2 \phi_2$$

The expected wage is increasing in x_2. When a lot of type 2 workers descend to island 1, it becomes easier for the ones who stayed in 2 to be the only applicant to a job and receive a high wage. When a type 2 worker looks for a job in island 1, he faces three possible outcomes: (a) with probability $e^{-q_2 x_2 h_2} e^{-q_1(1-x_1)}$, he is the only applicant and receives $\beta p_2 \phi_1$. Note that he produces less than in his "home" island and thus receives a lower wage than would have been the case if he had been the only applicant to a type 2 firm, (b) with probability $1-e^{-q_2 x_2 h_2}$, he is in competition with other type 2 workers and receives 0 (regardless of whether he ends up employed or unemployed), and (c) with probability $e^{-q_2 x_2 h_2}(1-e^{-q_1})$, he is in competition with type 1 workers only and receives $\phi_1(p_2-p_1)$.[19] The expected payoff of a worker type 2 who searches for a job in island 1, $\omega_{2,1}$, is thus

$$E\omega_{2,1} = \beta e^{-q_2 x_2 h_2} e^{-q_1} p_2 \phi_1 + \beta e^{-q_2 x_2 h_2} \phi_1 (p_2 - p_1) \left[1 - e^{-q_1}\right]$$

The expected wage in island 1 is decreasing in x_2: when there are fewer type 2 workers in island 1, there is less competition in island 1, and type 2 workers can expect a higher wage. In equilibrium, a type 2 worker must be indifferent between looking for a job in island 2 or in island 1. This arbitrage condition, $A(x_2)$, determines, x_2, the equilibrium allocation of workers

$$A(x_2) = -e^{-q_2(1-x_2)} p_2 \phi_2 + e^{-q_1 x_2 h_2} e^{-q_1} p_2 \phi_1 + e^{-q_1 x_2 h_2} \phi_1 (p_2 - p_1) \left[1 - e^{-q_1}\right] = 0 \quad (9)$$

Consider now the problem of type 1 workers. Type 1 workers could choose to move up the occupation ladder and search for a job in island 2. This will not happen as long as there are type 2 workers in island 1. The intuition is that, as long as type 2 workers are indifferent between their "home" island and the island below, type 1 workers will always prefer to remain in 1. Type 2 (high types) workers capture a higher share of the surplus from firms in island 1 than from firms in island 2, and yet they are indifferent between the two islands. Type 1 workers (low types), in contrast, capture the same share of the surplus in both islands. Thus, in equilibria with under-employment of the high-types (the equilibria we are interested in in this paper), low types strictly prefer to stay in island 1. Formally, the expected wage that an agent of type 1 would earn in island 2 is always lower than her expected wage in island 1. Indeed, $E\omega_{1,2} = e^{-q_2(1-x_2)} p_1 \phi_2$ and equation (9) implies $e^{-q_2(1-x_2)} p_2 \phi_2 < e^{-q_1 x_2 h_2} e^{-q_1} p_2 \phi_1$.

Unicity

Under condition (C_2^p), some workers of type 2 will always apply in island 1. We have already shown that, as long as type 2 workers are indifferent between the 2 islands, there cannot be workers of type 1 looking for jobs in island 2. As a consequence, the only variable that adjusts is the number of workers of type 2

[19] As noted earlier, despite the presence of competing applicants, a single type 2 applicant can extract some of the surplus thanks to due to his productivity advantage over the other applicants.

applying in island 1.

The trade-off faced by type 2 workers is monotonic, i.e. as they apply more in island 1, their relative gain of doing so is strictly decreasing. As already discussed above, under condition (C_2^p), the relative gain of applying in island 1 is initially positive (for $x_2 = 0$). The relative gain is negative when $x_2 = 1$, because $p_2\phi_2 > p_2\phi_1$. It only crosses once the x-axis, and the intersection defines the unique equilibrium. $\square$

Proof. Proposition 2

We directly consider the general problem of a worker $n \in \{1, \ldots, N\}$ who can decide to look for a job in his home island, or instead move down the occupation ladder to look for a job. In the spirit of the $N = 2$ case, we can exclude the possibility that workers look for jobs in higher technology islands or that they descend to lower levels than the one immediately below. The intuition is the same as the one developed in the 2 islands case: as long as a particular type is indifferent between two islands, the more (resp. less) skilled types will always prefer the island above (resp. below). The reason lies in the fact that the relative rent extracted between island $n-1$ and n is increasing in the skills of agents.

A type n worker has two choices, he can (i) look for a job in island n, his "home island", or (ii) look for a job in island -1, i.e., move down the occupation ladder. As in Proposition 1, we consider these two possibilities, and the only difference with the $N = 2$ case is that workers now have to take into account the fact that some higher type workers may be looking for work in their home island.

When a type n worker looks for a job in island n, he faces two possible outcomes: (a) with probability $e^{-q_n(1-x_n)}e^{-q_n x_{n+1} h_{n+1}}$, he is the only applicant and receives $\beta p_n \phi_n$, or (b), with probability $1 - e^{-q_n(1-x_n)}e^{-q_n x_{n+1} h_{n+1}}$, he is in competition with other workers (either from his own island n or from island $n+1$) and receives 0 (regardless of whether he ends up employed or unemployed). The expected payoff of a worker type n who searches for a job in island n, $\omega_{n,n}$, is thus

$$E\omega_{n,n} = \beta e^{-q_n(1-x_n)} e^{-q_n x_{n+1} h_{n+1}} p_n \phi_n$$

Consider now the case in which worker type n moves down to island $n-1$. There are 3 possibilities: (a) with probability $e^{-q_{n-1} x_n h_n} e^{-q_{n-1}(1-x_{n-1})}$, he is the only applicant and receives $\beta p_n \phi_{n-1}$, (b) with probability $1 - e^{-q_{n-1} x_n h_n}$, he is in competition with type n workers coming, like him, from the island above, and he receives 0 (regardless of whether he ends up employed or unemployed), and (c), with probability $e^{-q_{n-1} x_n h_n} \left(1 - e^{-q_{n-1}(1-x_{n-1})}\right)$, he is in competition with type $n-1$

workers only and receives $\phi_{n-1}(p_n - p_{n-1})$.[20] The expected payoff of a worker type n who searches for a job in island $n-1$, $\omega_{n,n-1}$, is thus

$$E\omega_{n,n-1} = \beta e^{-q_{n-1}x_n h_n} e^{-q_{n-1}(1-x_{n-1})} p_n \phi_{n-1} + \beta e^{-q_{n-1}x_n h_n} \left[1 - e^{-q_{n-1}(1-x_{n-1})}\right] \phi_{n-1}(p_n - p_{n-1})$$

In equilibrium, a type n worker must be indifferent between staying in island n or moving down to island $n-1$, which implies the arbitrage equation

$$\begin{aligned} A_n(x_{n+1}, x_n) &= -e^{-q_n(1-x_n)} e^{-q_n x_{n+1} h_{n+1}} p_n \phi_n \\ &+ e^{-q_{n-1}x_n h_n} e^{-q_{n-1}(1-x_{n-1})} p_n \phi_{n-1} + e^{-q_n x_n h_n} \left[1 - e^{-q_{n-1}x_n h_n}\right] = 0 \end{aligned}$$

These equations characterize the equilibrium allocation.

Unicity

As in the $N = 2$ case, uniqueness comes from a monotonicity argument.

Condition (C_3^p) implies that there will be some high-skilled workers descending even when all mid-skilled workers are applying in island 2. Condition (C_2^p) implies that mid-skilled workers descend even when none of the high-skilled workers are applying in their island. As in the case $N = 2$, some high skilled workers always apply in island 3 in island 3 because $p_3\phi_3 > p_3\phi_2$.

Under this set of conditions, we now show that there exists a unique equilibrium. With two types of actors, the relative gain depends on the others' behaviors: there is a complementarity between their choices. To see it, let us write down the conditions under which wages are equal for workers 3 in islands 2 and 3, and workers 2 in islands 1 and 2.

$$\begin{cases} p_3\phi_3 e^{-q_3(1-x_3)} = \phi_2 \left[p_3 - p_2 + p_2 e^{-q_2(1-x_2)}\right] e^{-q_2 x_3 h_3} & (A_3) \\ p_2\phi_2 e^{-q_2(1-x_2) - q_2 x_3 h_3} = \phi_1 \left[p_2 - p_1 + p_1 e^{-q_1}\right] e^{-q_1 x_2 h_2} & (A_2) \end{cases}$$

The two curves (A_3) and (A_2) both describe a positive relationship between x_3 and x_2, respectively the pull $x_3 = f_3(x_2)$ and push $x_3 = f_2(x_2)$ effects. Any interior equilibrium should be at the intersection of those two curves. It can be shown that:

$$\begin{cases} f_3'(x_2) = \frac{q_1 h_2 + q_2}{q_2 h_3} \\ f_2'(x_2) = \frac{q_2}{q_3 + q_2 h_3} \frac{p_2 e^{-q_2(1-x_2)}}{p_3 - p_2 + p_2 e^{-q_2(1-x_2)}} \end{cases}$$

It can be easily verified that $\frac{q_2}{q_3+q_2 h_3}\frac{p_2}{p_3} < \frac{q_1 h_2 + q_2}{q_2 h_3}$. As a consequence, (A_2) is always steeper than (A_3), e.g. the push effect is always stronger than the pull effect and uniqueness derives from this observation (see figure 9).

[20] As noted earlier, despite the presence of competing applicants, a single type n applicant can extract some of the surplus thanks to due to his productivity advantage over the other applicants.

□

Proof. Proposition 3

The workers no-arbitrage conditions are already derived in Proposition 1. The number of job openings is given by the free entry condition, and we only need to express the firm's expected profit as a function the the number of job openings V_1, or the "initial" queue length $q1 = \frac{L_1}{V_1}$.

Consider first a firm that enters island 1 and has therefore technology ϕ_1. The firm's profit will depend on the number of applications it receives. There are 5 cases: (the outcome of the wage negotiation process in each case is described in detail in the Proof of Proposition 1)

1. The firm has no applicant. Profit is zero.

2. The firm has only one applicant. The firm gets a share $1 - \beta$ of the output, i.e. $(1 - \beta)p_2\phi_1$ if the applicant is of type 1 (which happens with probability $P(a_1 = 1, a_2 = 0) = q_1 e^{-q_1} e^{-q_2 x_2 h_2}$), and $(1 - \beta)p_2\phi_1$ if the applicant is of type 2 (which happens with probability $P(a_1 = 0, a_2 = 1) = q_2 x_2 h_2 e^{-q_2 x_2 h_2} e^{-q_1}$).

3. The firm has more than one applicants of type 1 (and no applicants of type 2). The firm gets all the surplus: $p_1\phi_1$. This happens with probability $e^{-x_2 h_2 q_1}\left[1 - e^{-q_1} - q_1 e^{-q_1}\right]$.

4. The firm has more than one applicants of type 2 (and no applicants of type 1). The firm gets all the surplus: $p_2\phi_1$. This happens with probability $1 - e^{-x_2 h_2 q_1} - x_2 h_2 q_1 e^{-x_2 h_2 q_1}$.

5. The firm has more than one applicants *of different types*. The most productive worker is hired and gets a share β of the surplus generated over hiring the second-best applicant. The firm generates a profit $p_1\phi_1 + (1 - \beta)\phi_1(p_2 - p_1)$. This happens with probability $x_2 h_2 q_1 e^{-x_2 h_2 q_1}(1 - e^{-q_1})$.

The expected profit for a firm with technology 1 is then as stated in Proposition 3. Reasoning similarly gives the expected profit for a firm with technology 2. Consequently, free entry imposes two no-profit conditions in addition to workers' arbitrage equations described in Section 3.

The unicity of the equilibrium comes is a direct consequence of Corollaries 1 and 2 that we prove next. □

Proof. Corollary 1

First, it is straightforward from the expression of $\pi_2(x_2, q_2)$ that the free entry condition $\pi_2 = c_2$ imposes that $q_2(1 - x_2)$ is constant, so that the expected wage in island 2, $E\omega_{2,2}$, is constant. We can thus restrict our analysis to the arbitrage condition coupled with the free entry condition in island 1.

$$\begin{cases} E\omega_{2,2} = \phi_1 \left[p_2 - p_1 + p_1 e^{-q_1}\right] e^{-q_1 x_2 h_2} & (L^S) \\ p_2\phi_1 - c_1 = \phi_1 \left[(p_2 - p_1 + p_1 e^{-q_1})\left(1 + (2-\beta) q_1 x_2 h_2\right) + (2-\beta) p_1 q_1 e^{-q_1}\right] e^{-q_1 x_2 h_2} & (L_1^D) \end{cases}$$

The (L_1^D) equation defines a job creation function $q_1(x_2)$. As before, we only consider interior solutions, i.e. we impose that:

$$\phi_1 \left[p_2 - p_1 + p_1 e^{-q_1(1)}\right] e^{-q_1(1) h_2} < E\omega_{2,1} < \phi_1 \left[p_2 - p_1 + p_1 e^{-q_1(0)}\right]$$

Under this condition, the relative gain of searching for a job in lower-tech island is positive for $x_2 = 0$ ($E\omega_{2,2} < E\omega_{2,1}$) and negative for $x_2 = 1$ ($E\omega_{2,2} > E\omega_{2,1}$).

Combining the (L_1^D) and (L^S) equations, it can be shown with a little bit of algebra that:

$$\begin{aligned} E\omega_{2,1}^{'}(x_2) &= \tfrac{\partial E\omega_{2,1}}{\partial x_2} + q_1^{'}(x_2)\tfrac{\partial E\omega_{2,1}}{\partial q_1} \\ &= \tfrac{(2-\beta) q_1 (E\omega_{2,1} - E\omega_1)}{[(2-\beta) q_1 x_2 h_2 - (1-\beta)] E\omega_{2,1} + (2-\beta) q_1 E\omega_1} q_1^{'}(x_2) E\omega_1 < 0 \end{aligned}$$

and that $q_1^{'}(x_2)\frac{\partial E\omega_{2,1}}{\partial q_1} > 0$.

This proves Corollary 1. Moreover, using that $E\omega_{2,1}(x_2) < 0$ with the fact that $E\omega_{2,2}(x_2)$ is constant guarantees the uniqueness of the equilibrium. □

Proof. Corollary 2

Combining the (L_1^D) and (L^S) equations, it can be shown with a little bit of algebra that:

$$\begin{aligned} E\omega_{1,1}^{'}(x_2, q_1(x_2)) &= \tfrac{\partial E\omega_1}{\partial x_2} + q_1^{'}(x_2)\tfrac{\partial E\omega_1}{\partial q_1} \\ &= \tfrac{[(2-\beta) q_1 x_2 h_2 - (1-\beta)](E\omega_1 - E\omega_{2,1})}{[(2-\beta) q_1 x_2 h_2 - (1-\beta)] E\omega_{2,1} + (2-\beta) q_1 E\omega_1} q_1^{'}(x_2) E\omega_1 \gtrless 0 \end{aligned}$$

We can see that $E\omega_{1,1}(x_2, q_1(x_2))$ is not monotonically decreasing, implying that a larger number of high-skilled workers does not necessarily imply lower expected income for low-skilled workers. For $\beta < 1$, $E\omega_{1,1}(x_2, q_1(x_2))$ is initially decreasing and then increases once $(2-\beta) q_1 x_2 h_2 > (1-\beta)$. This proves Corollary 2. □

Proof. Proposition 4

The proof for the $N = 3$ case is very similar to the $N = 2$ case. The equilibrium is characterized by the allocation of workers of type 3 and 2, and the free-entry conditions in islands 1, 2 and 3. First, the free entry condition imposes that $q_3(1-x_3)$ is constant, and thus the expected wage in island 3, $E\omega_{3,3}$, is constant. We can thus restrict our analysis to the arbitrage conditions for workers of type 2 and 3 coupled with the free entry conditions in island 1 and 2. □

Proof. Proposition 5

The maximization program of the central planner can be written as follows (denote Y the aggregate output of the economy):

$$\max_{x_2,q_1,q_2} \{Y\}$$

subject to

$$\begin{cases} \pi_2(x_2,q_2) = c_2 \\ \pi_1(x_2,q_1) = c_1 \end{cases}$$

We can already simplify the program through two channels. First, with free entry, the aggregate profit of firms (net of investment costs) is zero. Consequently, the central planner equivalently maximizes the wage bill of workers. Second, free entry in island 2 imposes that q_2 is set such as to make $(1-x_2)q_2$ constant.

$$(1-x_2)q_2 = f^{-1}\left(\frac{c_2}{p_2\phi_2}\right)$$

The program then sums up to:

$$\max_{x_2,q_1} \left\{(1-x_2)h_2E\omega_{2,2} + h_2x_2e^{-q_1h_2x_2}\phi_1\left[p_2-p_1+p_1e^{-q_1}\right] + e^{-q_1h_2x_2-q_1}p_1\phi_1\right\}$$

subject to

$$p_2\phi_1 - \phi_1e^{-x_2h_2q_1}\left[\left(p_2-p_1+p_1e^{-q_1}\right)(1+(2-\beta)x_2h_2q_1) + (2-\beta)p_1q_1e^{-q_1}\right] = c_1$$

The first-order condition in x_2 leads to:

$$A(x_2,q_1) - B_{x_2}(x_2,q_1) - \lambda C_{x_2}(x_2,q_1) = 0$$

where

$$\begin{cases} B_{x_2}(x_2,q_1) = q_1h_2\phi_1e^{-q_1h_2x_2}\left[x_2h_2\left(p_2-p_1+p_1e^{-q_1}\right)+p_1\phi_1e^{-q_1}\right] \\ C_{x_2}(x_2,q_1) = q_1h_2\phi_1e^{-q_1h_2x_2}\left[\left((2-\beta)q_1x_2h_2-(1-\beta)\right)\left(p_2-p_1+p_1e^{-q_1}\right)+(2-\beta)p_1\phi_1e^{-q_1}\right] \end{cases}$$

We can already see that profits losses are not entirely internalized by workers: the marginal gain in wages for workers of type 2 cannot fully translate in marginal profits for firms in island 1. The first-order condition in q_1 gives:

$$-B_{q_1}(x_2, q_1) - \lambda C_{q_1}(x_2, q_1) = 0$$

where

$$\begin{cases} B_{q_1}(x_2, q_1) = x_2/q_1 B_{x_2}(x_2, q_1) + (1 + x_2 h_2)\phi_1 p_1 e^{-q_1 x_2 h_2 - q_1} \\ C_{q_1}(x_2, q_1) = x_2/q_1 C_{x_2}(x_2, q_1) + [q_1(2-\beta)(1 + x_2 h_2) - (1-\beta)]\, \phi_1 p_1 e^{-q_1 x_2 h_2 - q_1} \end{cases}$$

We can observe the symmetry between the expressions in x_2 and q_1. The basic intuition is that, since profits can be written as a function of $q_1 x_2$, the externality generated by a change in x_2 will be partly compensated by an inverse change in q_1. Indeed, combining these two equations, we have that:

$$A(x_2, q_1) = \frac{B_{x_2}(x_2, q_1) C_{q_1}(x_2, q_1) - B_{q_1}(x_2, q_1) C_{x_2}(x_2, q_1)}{C_{q_1}(x_2, q_1)}$$

Once accounted for the expression of $B_{q_1}(x_2, q_1)$ and $C_{q_1}(x_2, q_1)$,

$$A(x_2, q_1) = \frac{(1-\beta)\, h_2 q_1 \phi_1^2 p_1 e^{-2q_1 x_2 h_2 - q_1}\, (p_2 - p_1)}{C_{q_1}(x_2, q_1)}$$

As a consequence, $A(x_2, q_1)$ is striclty positive as long as the surplus is not entirely given to workers, i.e. $\beta < 1$, and workers are not equally productive, i.e. $p_2 - p_1 > 0$. Coupled with the two free entry conditions, this equation characterizes the constrained optimum which does not coincide with the decentralized allocation. $A(x_2, q_1) > 0$ implies that wages are higher in 1 than in 2. In other words, the decentralized allocation induces a lower x_2, a higher q_1 and a lower q_2. $\square$

Proof. Proposition 6

We proceed here exactly as we did for Proposition 4. The maximization program of the central planner can be written as follows (denote Y the aggregate output of the economy):

$$\max_{x_2, x_3, q_1, q_2, q_3} \{Y\}$$

subject to

$$\begin{cases} \pi_3(x_3, q_3) = c_3 \\ \pi_2(x_3, x_2, q_2) = c_2 \\ \pi_1(x_2, q_1) = c_1 \end{cases}$$

As before, two remarks help us simplify the program. First, with free entry, the aggregate profit of firms (net of investment costs) is zero: the central planner maximizes the wage bill of workers. Second, free entry in island 3 imposes that q_3 is set such as to make $(1-x_3)q_3$ constant.

$$(1-x_3)q_3 = f^{-1}\left(\frac{c_3}{p_3\phi_3}\right)$$

The program then sums up to (where each line represents wages earn by agents of different types):

$$\max_{x_2,x_3,q_1,q_2}\left\{\begin{array}{l} h_2h_3(1-x_3)E\omega_{3,3} + h_2h_3x_3e^{-q_2h_3x_3}\phi_2\left[p_3 - p_2 + p_2e^{-q_2(1-x_2)}\right] \\ +h_2(1-x_2)p_2\phi_2e^{-q_2x_3h_3-q_2(1-x_2)} + h_2x_2\phi_1\left(p_2 - p_1 + p_1e^{-q_1}\right)e^{-q_1x_2h_2} \\ +p_1\phi_1e^{-q_1x_2h_2-q_1} \end{array}\right\}$$

subject to

$$\left\{\begin{array}{l} p_3\phi_2 - \phi_2e^{-x_3h_3q_2}\left[\left(p_3 - p_2 + p_2e^{-q_2(1-x_2)}\right)(1+(2-\beta)x_3h_3q_2) + (2-\beta)p_2q_2(1-x_2)e^{-q_2(1-x_2)}\right] = c_2 \\ p_2\phi_1 - \phi_1e^{-x_2h_2q_1}\left[\left(p_2 - p_1 + p_1e^{-q_1}\right)(1+(2-\beta)x_2h_2q_1) + (2-\beta)p_1q_1e^{-q_1}\right] = c_1 \end{array}\right.$$

We need now to write the four first-order conditions:

$$\left\{\begin{array}{ll} A_3(x_3,x_2,q_2) - B_{x_3}(x_3,x_2,q_2) - \lambda_2C_{x_3}(x_3,x_2,q_2) = 0 & [x_3] \\ -B_{q_2}(x_3,x_2,q_2) + \lambda_2C_{q_2}(x_3,x_2,q_2) = 0 & [q_2] \\ A_2(x_3,x_2,q_2,q_1) - B_{x_2}(x_3,x_2,q_2,q_1) - \lambda_2C_{x_2}(x_3,x_2,q_2) - \lambda_1D_{x_2}(x_3,x_2,q_1) = 0 & [x_2] \\ -B_{q_1}(x_2,q_1) - \lambda_2D_{q_1}(x_2,q_1) = 0 & [q_1] \end{array}\right.$$

Let us detail the notations, A_3 (resp. A_2) denotes the difference between wages earned in level 2 and 3 (resp. 1 and 2) for workers of type 3 (resp. 2). B_{x_3} and B_{q_2} represents the additional terms deriving from differentiating W with respect to x_3 and q_2. We report their exact expression below.

$$\left\{\begin{array}{ll} B_{x_3}(x_3,x_2,q_2) = & h_3q_2\phi_2e^{-x_3q_2h_3}\left[\left(p_3 - p_2 + p_2e^{-q_2(1-x_2)}\right)(1+x_3h_3q_2) + p_2q_2(1-x_2)e^{-q_2(1-x_2)}\right] \\ B_{q_2}(x_3,x_2,q_2) = & \frac{x_3}{q_2}B_{x_3}(x_3,x_2,q_2) + (x_3h_3+1-x_2)\,h_2\,(1-x_2)\,p_2\phi_2e^{-q_2h_3x_3-q_2(1-x_2)} \end{array}\right.$$

B_{x_2} and B_{q_2} represents the additional terms deriving from differentiating W with respect to x_3 and x_2. We report their exact expression below.

$$\left\{\begin{array}{ll} B_{x_2}(x_3,x_2,q_2,q_1) = & -q_2h_2p_2\phi_2e^{-q_2h_3x_3-q_2(1-x_2)}(x_3h_3+1-x_2) \\ & +q_1h_2\phi_1e^{-q_1h_2x_2}\left[h_2x_2(p_2-p_1+p_1e^{-q_1}) + p_1e^{-q_1}\right] \\ B_{q_1}(x_2,q_1) = & x_2h_2\phi_1e^{-q_1h_2x_2}\left[h_2x_2(p_2-p_1+p_1e^{-q_1}) + p_1e^{-q_1}\right] \\ & +(1+x_2h_2)p_1\phi_1e^{-q_1h_2x_2-q_1} \end{array}\right.$$

C_l represents the additional terms deriving from differentiating the profits in island 2 with respect to l. D_l represents the additional terms deriving from differentiating the profits in island 1 with respect to l. We report their exact expression below.

$$\begin{cases} C_{x_3}(x_3,x_2,q_2) = & h_3q_2\phi_2 e^{-x_3h_3q_2}\left[\left(p_3-p_2+p_2e^{-q_2(1-x_2)}\right)((2-\beta)x_3h_3q_2-(1-\beta))\right. \\ & \left.+(2-\beta)p_2q_2(1-x_2)e^{-q_2(1-x_2)}\right] \\ C_{q_2}(x_3,x_2,q_2) = & \frac{x_3}{q_2}C_{x_3}(x_3,x_2,q_2)+p_2\phi_2(1-x_2)e^{-x_3h_3q_2-q_2(1-x_2)}\left[(x_3h_3+1-x_2)(2-\beta)q_2-(1-\beta)\right] \\ D_{x_2}(x_3,x_2,q_1) = & h_2q_1\phi_1e^{-x_2h_2q_1}\left[(p_2-p_1+p_1e^{-q_1})((2-\beta)x_2h_2q_1-(1-\beta))\right. \\ & \left.+(2-\beta)p_1q_1e^{-q_1}\right] \\ D_{q_1}(x_2,q_1) = & \frac{x_2}{q_1}D_{x_2}(x_3,x_2,q_1)+p_1\phi_1e^{-x_2h_2q_1-q_1}\left[(x_2h_2+1)(2-\beta)q_1-(1-\beta)\right] \\ C_{x_2}(x_3,x_2,q_2) = & -q_2\phi_2p_2\left[(x_3h_3+1-x_2)(2-\beta)q_2-(1-\beta)\right]e^{-x_3h_3q_2-q_2(1-x_2)} \end{cases}$$

The main difference with the $N=2$ case comes from an additional interaction term between workers of type 2 and 3. Workers of type 2 influences the profits that firms can make in island 2. C_{x_2} represents this gain in profits.

We eliminate the shadow prices in the first-order conditions:

$$\begin{cases} A_3(x_3,x_2,q_2) = B_{x_3}(x_3,x_2,q_2) - \frac{C_{x_3}(x_3,x_2,q_2)}{C_{q_2}(x_3,x_2,q_2)}B_{q_2}(x_3,x_2,q_2) \\ A_2(x_3,x_2,q_2,q_1) = B_{x_2}(x_3,x_2,q_2,q_1) - \frac{C_{x_2}(x_3,x_2,q_2)}{C_{q_2}(x_3,x_2,q_2)}B_{q_2}(x_3,x_2,q_2) - \frac{D_{x_2}(x_3,x_2,q_1)}{D_{q_1}(x_2,q_1)}B_{q_1}(x_2,q_1) \end{cases}$$

Let us focus on the first equation:

$$A_3(x_3,x_2,q_2) = \frac{B_{x_3}(x_3,x_2,q_2)C_{q_2}(x_3,x_2,q_2) - C_{x_3}(x_3,x_2,q_2)B_{q_2}(x_3,x_2,q_2)}{C_{q_2}(x_3,x_2,q_2)}$$

As in proposition 4,

$$A_3(x_3,x_2,q_2) = \frac{(1-\beta)\,h_3h_2(1-x_2)^2q_2\phi_2^2p_2e^{-2q_2x_3h_3-q_2(1-x_2)}\,(p_3-p_2)}{C_{q_2}(x_3,x_2,q_2)}$$

It is easy to see that $A_3(x_3,x_2,q_2)>0$. The centralized allocation gives a higher wage to agents 3 in island 2 than what they would receive in island 3. x_3 is lower than in the decentralized allocation, q_3 is higher.

We now turn to the second equation.

$$A_2(x_3,x_2,q_2,q_1) = \frac{B_{x_2}D_{q_1}-D_{x_2}B_{q_1}}{D_{q_1}} - \frac{C_{x_2}}{C_{q_2}}B_{q_2}$$

$$\begin{aligned} A(x_2,q_1) = & \frac{(1-\beta)h_2q_1\phi_1^2p_1e^{-2q_1x_2h_2-q_1}(p_2-p_1)}{D_{q_1}(x_2,q_1)} \\ & +\frac{(1-\beta)(1-x_2)(p_3-p_2)\phi_2^2h_2p_2q_2x_3h_3e^{-2q_2x_3h_3-q_2(1-x_2)}}{C_{q_2}} \end{aligned}$$

□

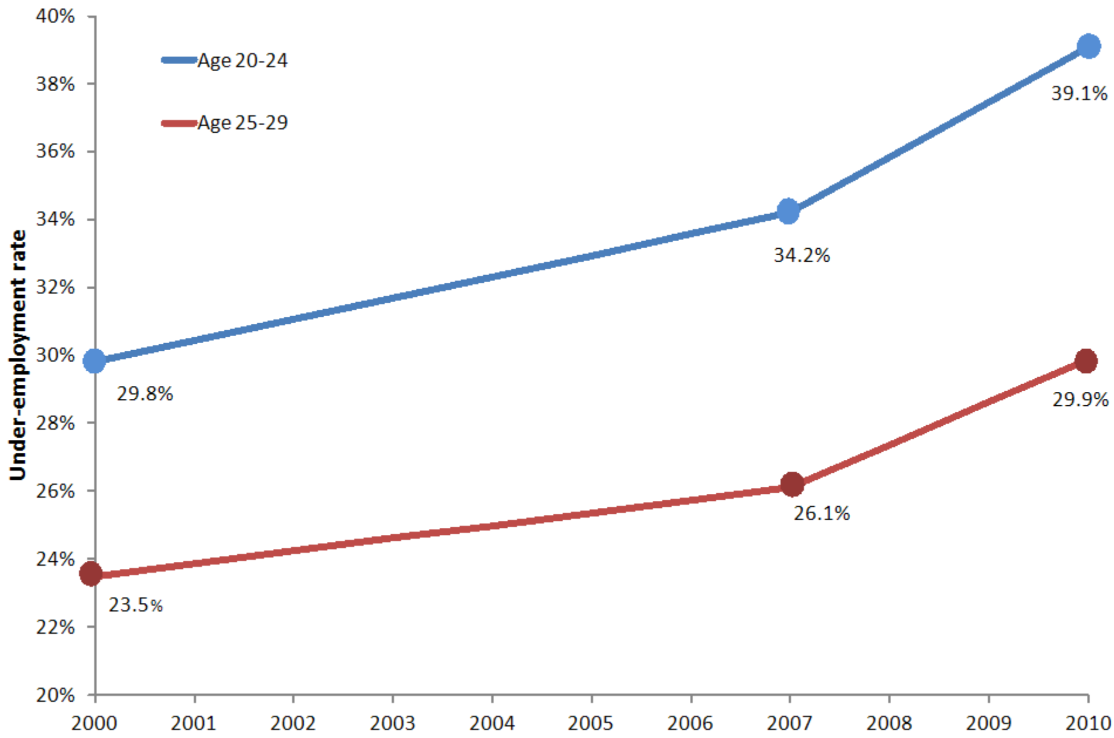

Figure 1. Share of graduates in occupations with lower required education.

Source: The State of Working America, 12th Edition (Lawrence Mishel, Josh Bivens, Elise Gould, Heidi Shierholz, Economic Policy Institute).

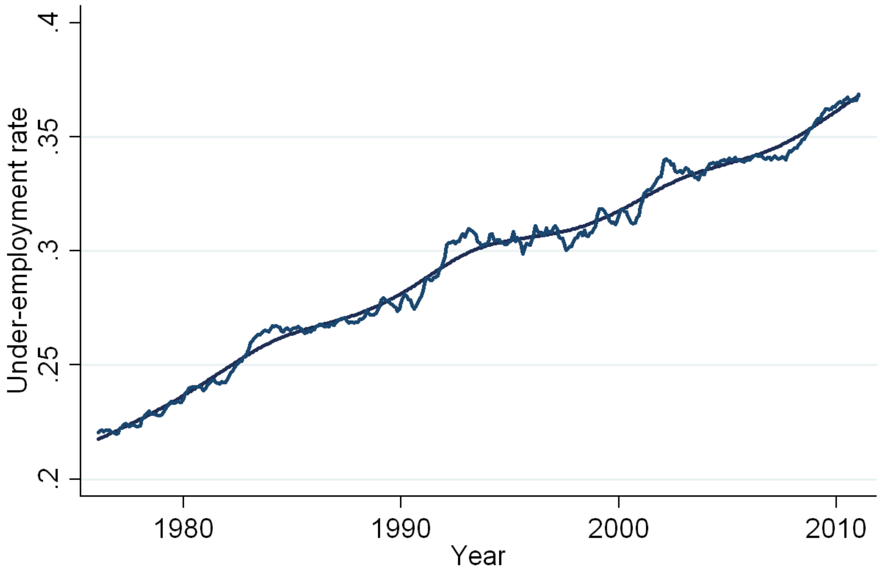

Figure 2. Under-employment: fraction of new hires, in which the hired candidate is more educated than required for the position.

Source: Current Population Survey, 1976-2012. Under-employment is defined as the fraction of hires in each month for which the *required* educational attainment is lower than the attainment of the hired worker (we consider 3 broad categories: less than high school, bachelor and post-graduates).

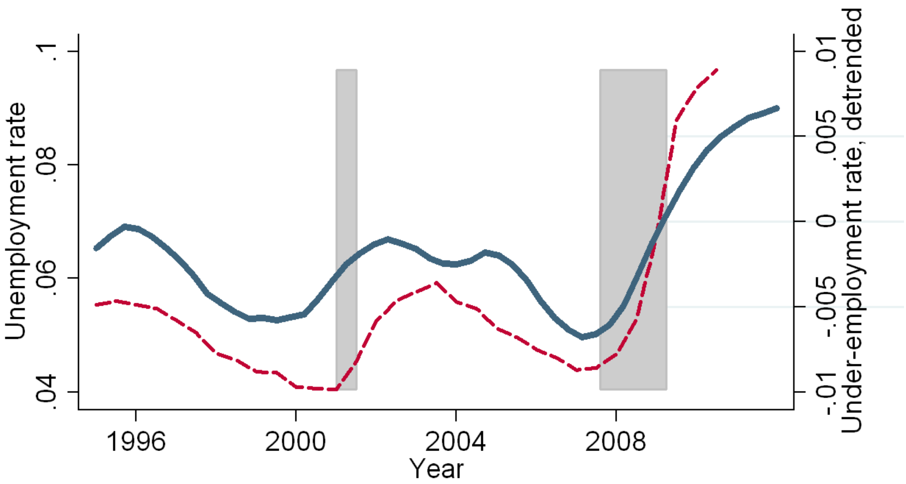

Figure 3. Detrended under-employment (blue) and unemployment (red).

Source: Current Population Survey, 1994-2012. Under-employment is defined as the fraction of hires in each month for which the required educational attainment is lower than the attainment of the hired worker (we consider 3 broad categories: less than high school, bachelor and post-graduates). Under-employment is detrended (linear trend). Both series are smoothed with a Moving Average (window of 12 months).

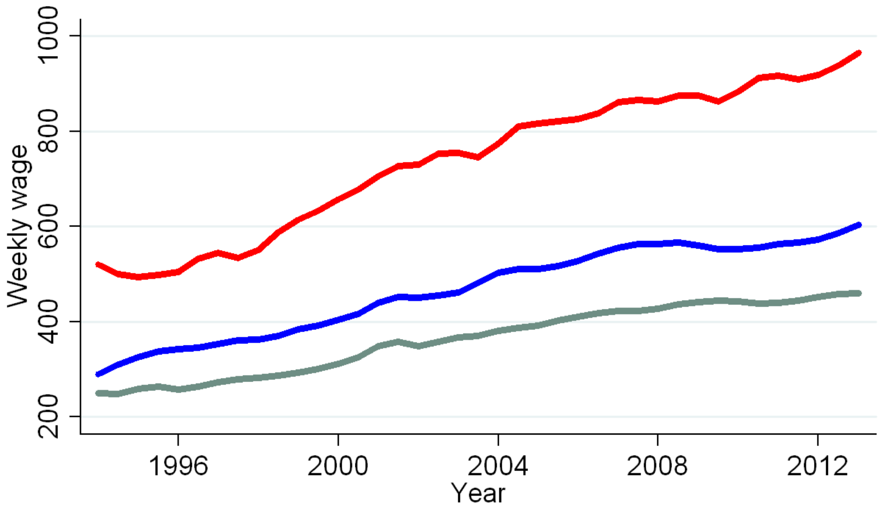

Figure 4. Hiring wages for (i) bachelors in jobs requiring a bachelor degree (red), in jobs with low requirements (blue), and (ii) non-bachelors in jobs with low requirements (teal).

Source: Current Population Survey, 1994-2012. Wages are here the average weekly nominal wages of new hires. All series are smoothed with a Moving Average (window of 12 months). Over the period, the wage premium of holding a bachelor for lower requirement jobs is 22%. This premium decreases to 14% once controlled for age, state dummies and occupation dummies. Over the period, the premium of working in high-requirement jobs for graduates is 35%. This premium decreases to 28% once controlled for age and state dummies.

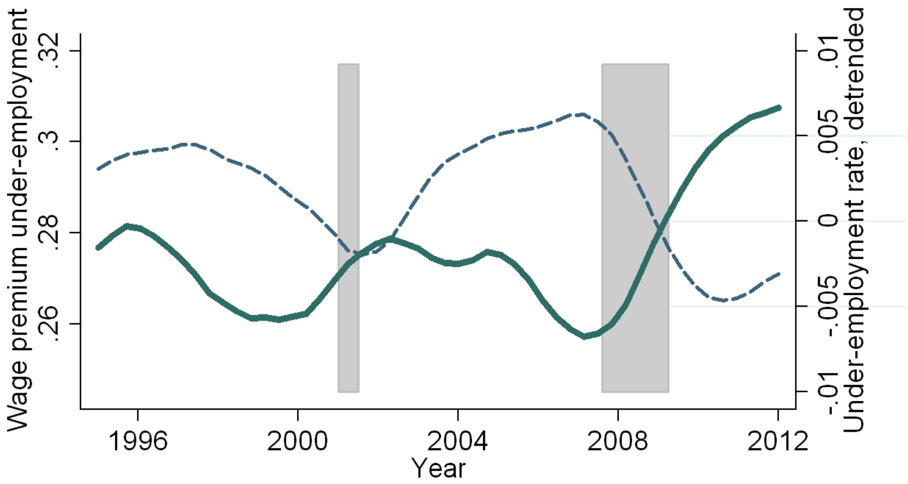

Figure 5. Education premium (plain line) for graduates hired in jobs with low requirements versus non-graduates and under-employment rate (detrended, dashed line).

Source: Current Population Survey, 1994-2012. The wage premium is $\omega_{2,1}/\omega_{1,1} - 1$ where $\omega_{2,1}$ (resp. $\omega_{1,1}$) is the average weekly nominal wage of new hires (resp. not) holding a bachelor degree in occupations requiring less than high school. Under-employment is detrended (linear trend). Both series are smoothed with a Moving Average (window of 12 months).

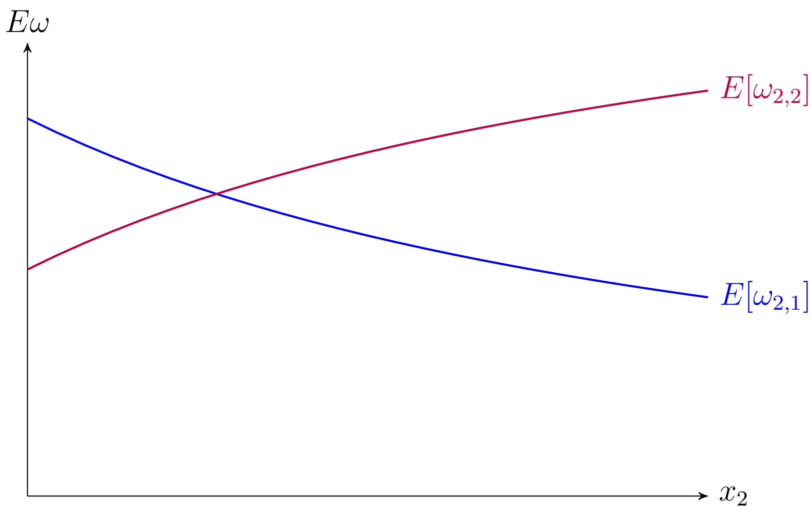

Figure 6. Partial Equilibrium – N=2.

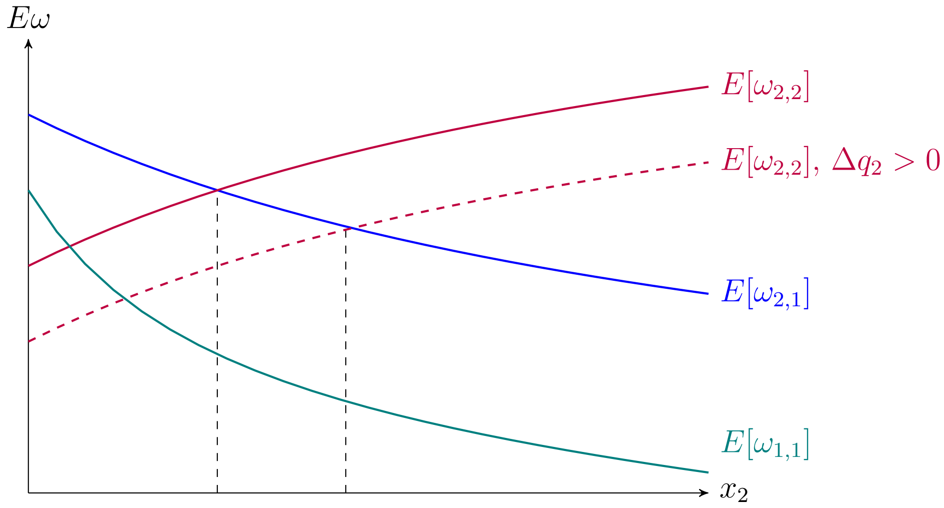

Figure 7. Partial Equilibrium – N=2 – shock on $\Delta q_2 > 0$.

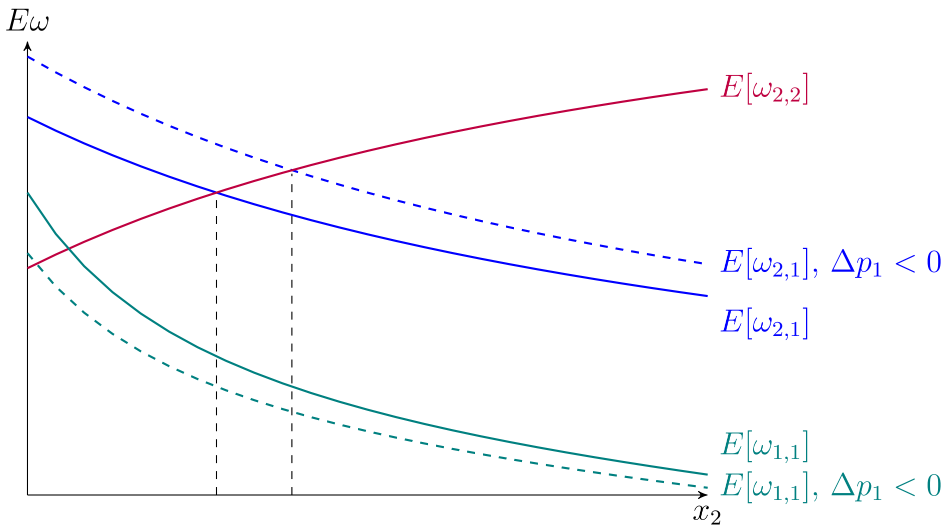

Figure 8. Partial Equilibrium – N=2 – shock on $\Delta p_1 < 0$.

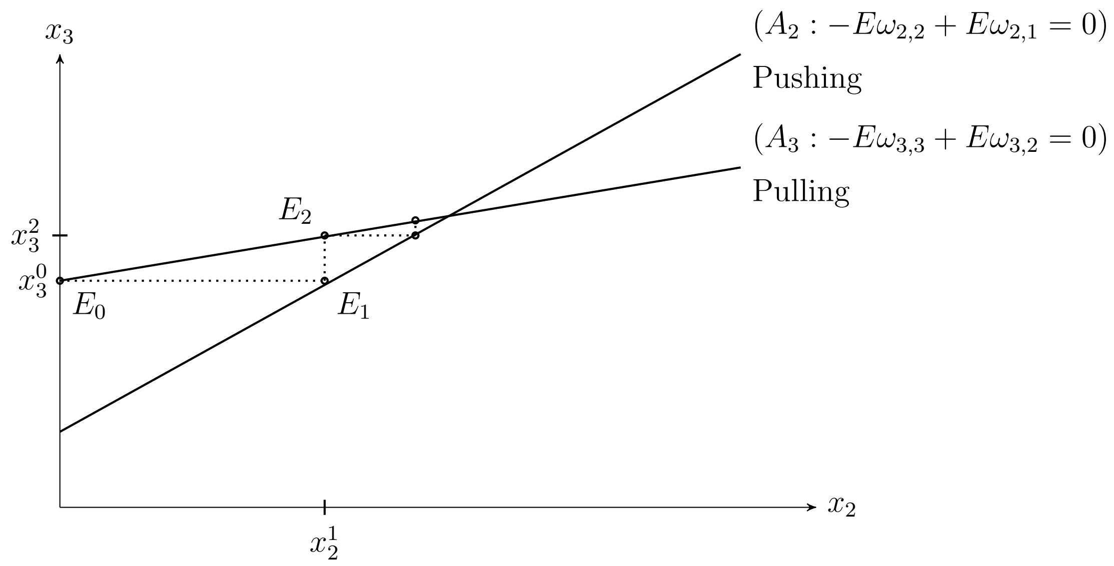

Figure 9. Partial Equilibrium – N=3.

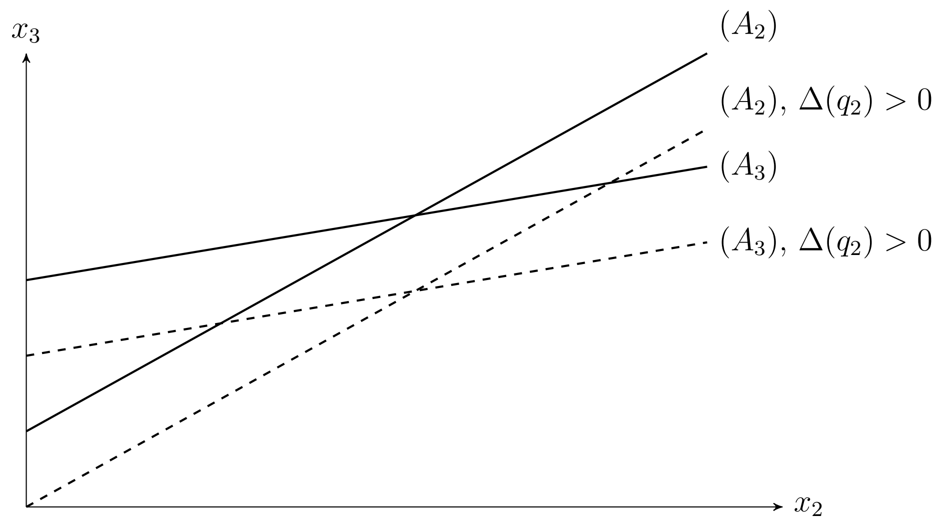

Figure 10. Partial Equilibrium – Effect of job polarization – N=3.

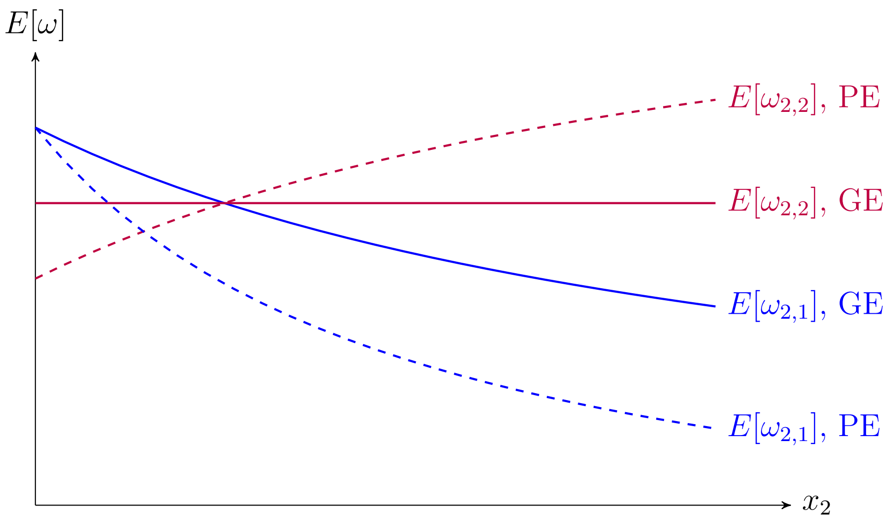

Figure 11. Expected wages for high-skilled workers (GE) – N=2.

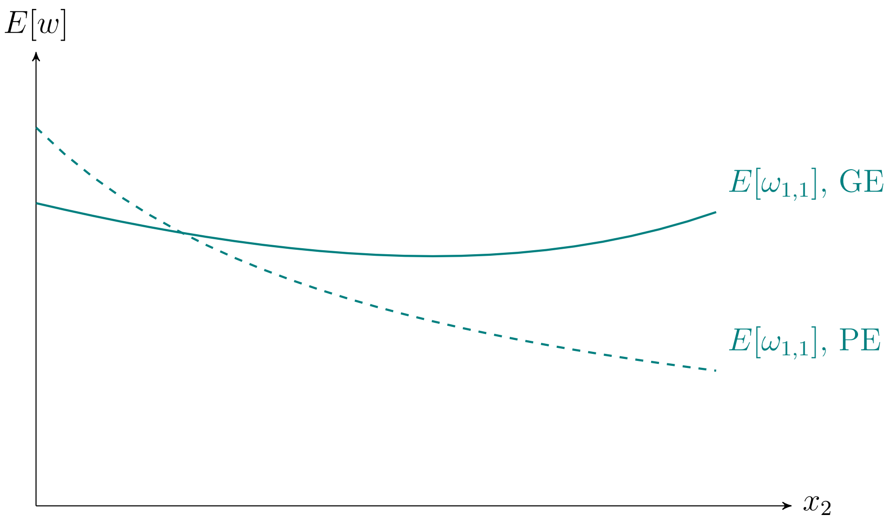

Figure 12. Expected wages for low-skilled workers (GE) – N=2.

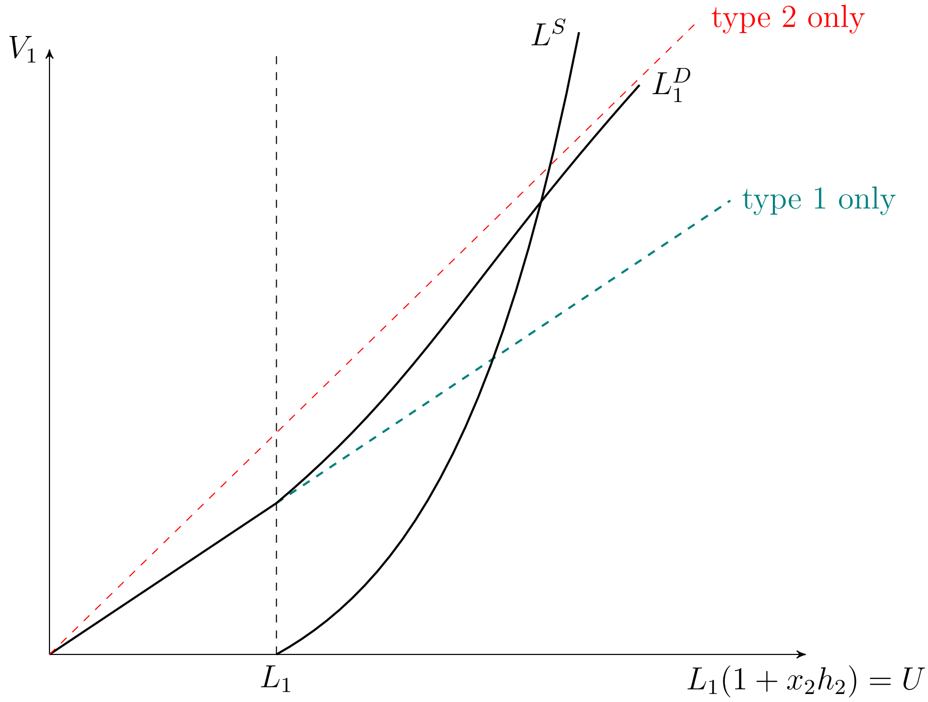

Figure 13. Labor market equilibrium – N=2.

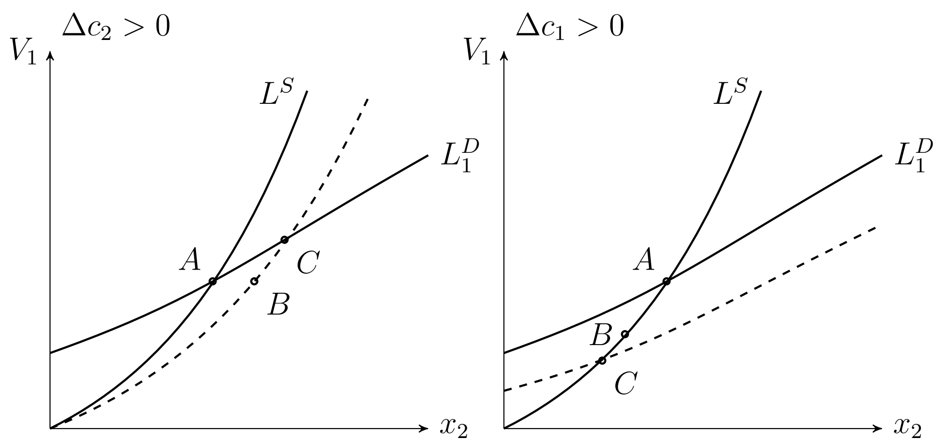

Figure 14. General Equilibrium – Comparative statics – N=2.

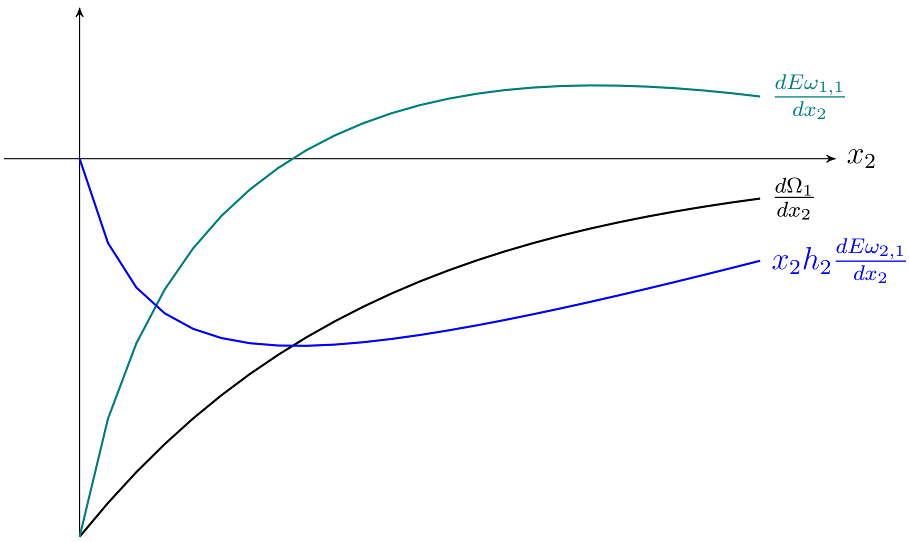

Figure 15. Inefficiency – ranking externality $\frac{d\Omega_1}{dx_2} = \frac{dE\omega_{1,1}}{dx_2} + x_2h_2\frac{dE\omega_{2,1}}{dx_2}$.

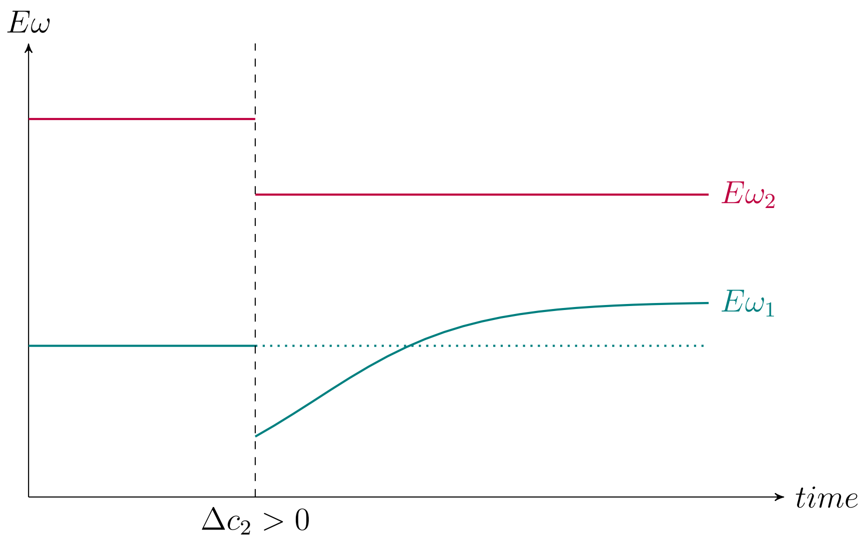

Figure 16. Shock (permanent loss of vacancies in island 2) – N=2.

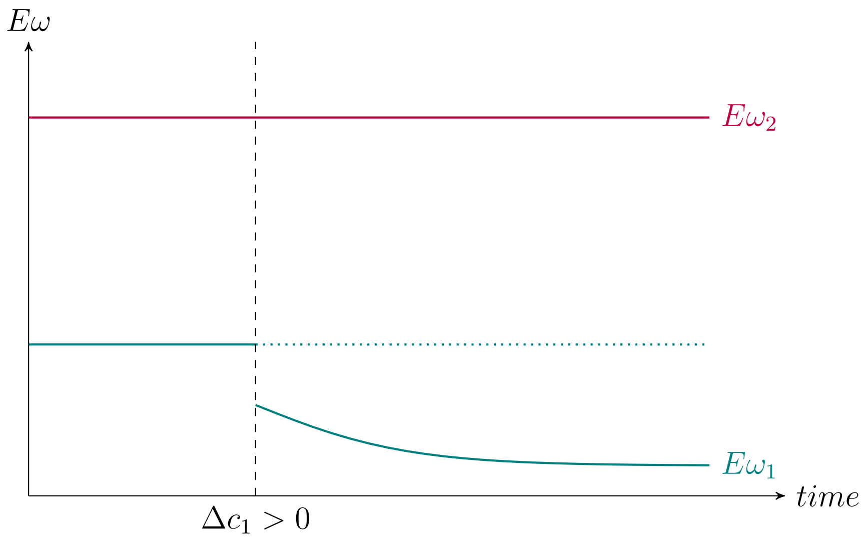

Figure 17. Shock (permanent loss of vacancies in island 1) – N=2.