Under-Employment and the Trickle-Down of Unemployment*

Regis Barnichon   Yanos Zylberberg

July 19, 2016

Abstract

A substantial fraction of workers are under-employed, i.e., employed in jobs for which they are over-qualified, and that fraction is strongly counter-cyclical. To explain these facts, we propose a search model in which a vacancy can simultaneously receive multiple applications. Through the wage-bargaining process, the model endogenously generates a "ranking" mechanism, in which high-skill applicants are systematically hired over less-skilled competing applicants. In equilibrium, some high-skill job seekers become under-employed to escape the competition for high-skill jobs and find a job more easily. In a dynamic model, an adverse aggregate shock increases under-employment, as high-skill job seekers escape the increased competition for high-skill jobs by moving down the job-ladder in greater proportion.

JEL: E24, J22, J64.

*Barnichon: CREI, Universitat Pompeu Fabra, CEPR; rbarnichon@crei.cat; Zylberberg: Bristol University; yanos.zylberberg@bristol.ac.uk. We thank Paula Bustos, Cynthia Doniger, Juan Dolado, Jan Eeckhout, Manolis Galenianos, Jordi Gali, Nicola Gennaioli, Marcel Jansen, Philipp Kircher, Alberto Martin, Guido Menzio, Matt Notowidigdo, Michael Reiter, Thijs van Rens, Robert Shimer, Jaume Ventura, Montserrat Vilalta-Bufi, Ludo Visschers, Till von Wachter and seminar participants at Bristol University, CERGE-EI, CREI, EUI, IAE, Kiel Institut für Weltwirtschaft, Royal Holloway, Universität Bonn, Universitat de Barcelona, Université de Lausanne, Universitat Pompeu Fabra, the 2015 Warwick Mismatch Workshop, the 16th CEPR/IZA European Summer Symposium in Labour Economics, the 10th Joint ECB/CEPR Labour Market Workshop for useful comments. We also thank Matt Notowidigdo for sharing his experimental data on employers' call-back rates. Barnichon acknowledges financial support from the Spanish Ministerio de Economia y Competitividad (grant ECO2011-23188), the Generalitat de Catalunya (grant 2009SGR1157) and the Barcelona GSE Research Network.

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

While the unemployment rate is the traditional gauge of the labor market, this paper argues that it misses an important dimension of the state of the labor market: under-employment. In the US, a substantial fraction of workers are under-employed, i.e., employed in jobs for which they are over-qualified, and that fraction –the under-employment rate– is strongly counter-cyclical, increasing markedly in slack labor markets. As shown in figure 1, the fraction of US college graduates working in lower skill-requirement occupations increased from 38.5 percent in 2008 to 41.5 percent in 2012. In other words, in those four years, the number of under-employed workers increased by 3 million; almost half as much as the increase in unemployed workers (about 7 million workers) over the same time period.

While under-employment finds a large echo in the media, there is surprisingly little work on the determinants of under-employment and its implications for business cycle fluctuations. In this paper, we study the characteristics and the determinants of under-employment. First, we document new stylized facts about under-employment, most notably its counter-cyclicality. Second, we provide a model to explain the existence and counter-cyclicality of under-employment. The model shows how fluctuations in under-employment can be a regressive redistributive mechanism, as higher-skill workers use under-employment to smooth labor demand shocks at the expense of lower-skill workers.

We start by studying empirically the phenomenon of under-employment in the US and show three stylized facts using CPS micro data: (i) under-employment is strongly counter-cyclical, (ii) under-employment is costly, an under-employed worker earning about 30 percent less than his non-under-employed counterpart, and (iii) under-employment is a persistent state with more than 70 percent of newly- under-employed workers still under-employed one year later.

Then, we aim to understand the economic mechanisms behind the existence and cyclicality of under-employment. We argue that under-employment is hard to rationalize with modern labor market models based on search frictions, be it a search and matching model à la Pissarides (1985) or a competitive search model à la Moen (1997). In those models, given the wage loss associated with under-employment, under-employment can only arise in equilibrium if low-qualification jobs are relatively more abundant than high-qualification jobs. However, this condition does not hold in the data: finding a job in a low-qualification occupation takes, on average,

[1]The Wall Street Journal, 26 March 2013

just as long as finding a job in a high-qualification occupation.

We argue that modern search models cannot explain the existence (and cyclicality) of under-employment, because they rely on the assumption that matching is random, in the sense that any two workers searching in the *same* labor market (and with the same search intensity) are equally likely to form a match. With random matching, high-skill workers find low-qualification jobs at the same rate as low-skill workers, giving them little incentive to move down the job ladder given the associated wage loss.

Instead, we propose a model with an alternative matching process, in which a vacancy can simultaneously receive multiple applications. The labor market is segmented into distinct islands with different productivity levels. Each worker, characterized by his skill level, can direct his search to a given island, which 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 applications, hiring is not random: applicants compete for the job during wage bargaining, and the firm ends up hiring the most profitable applicant. The negotiated wage then depends on both the number and the skill of other applicants.

Through the wage bargaining process, the model generates endogenously a "ranking" mechanism favoring high-skill job seekers, as high-skill applicants are systematically hired over less-skilled competing applicants. In this framework, under-employment (i) exists in equilibrium and (ii) is counter-cyclical. First, with ranking, under-employment exists, not because low-qualification jobs are more abundant, but because the competition to get a low-qualification job is, from the viewpoint of high-skill workers, less intense. Intuitively, a high-skill worker moves down the occupational ladder, in order to find a job more easily and escape competition from his high-skill peers. Second, with ranking, under-employment is counter-cyclical. Intuitively, following an adverse aggregate labor demand shock affecting all types of jobs, high-skill workers are less affected by the drop in low-skill jobs because of their ranking advantage. As a result, they smooth the aggregate shock by moving down the job ladder in greater proportion, and under-employment increases in recessions.

This mechanism in turn implies that under-employment has distributional implications. When high-skill workers move down the job ladder, they partially jump the unemployment queue, because they are hired over less-skilled competing applicants. But in turn, these these less-skilled workers are driven out of their market and further down the occupational ladder. Through this process, unemployment trickles down from the upper-occupation groups to the lower-occupation groups. Over the business cycle, high-skill workers smooth shocks at the expense of lower-skill workers by adjusting their under-employment rate, and in doing so they exacerbate the income volatility faced by lower-skill workers.

We show that under-employment is generally inefficient. Although the constrained optimal allocation in which the planner allocates job seekers to islands may call for some level of under-employment in order to maximize the matching probability of the most-skilled workers, the low-productivity island is too attractive to high-skill workers and, and there is too much under-employment in the decentralized allocation. As a result, compared to the constrained optimal allocation, too many high-skill workers are lost working for low productivity firms, and too many low-skill workers end up unemployed.

In a final section, we take a dynamic stochastic general equilibrium version of our model to the data, and we find that a calibrated model with aggregate productivity shocks alone can generate counter-cyclical movements in under-employment that are in line with the US experience. We also find that under-employment can be a powerful redistributive mechanism in which aggregate shocks get redistributed from the high skills to the low skills: In a simulation with aggregate shocks alone, the value of unemployment for high-skill workers is about 22 percent *less* volatile than the value of unemployment for loss-skill workers (despite suffering the same aggregate shocks).

The phenomenon of under-employment relates closely to the phenomenon of over-education,[2] which 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 (Freeman, 1976; McGuinness, 2006).[3] While the literature on over-education has focused on low-frequency patterns, our interest is instead on business cycle fluctuations. While the counter-cyclical nature of under-employment has, as far as we know, not been documented before, our finding echoes an old literature on "cyclical upgrading"; the possibility that the quality of matches improve in tighter labor markets (Reder, 1955; Okun, 1973).

An important line of research explains the existence of over-education with career mobility and on-the-job search (Sicherman, 1991; Sicherman and Galor, 1990; Dolado, Jansen and Jimeno, 2009), as high-educated workers can choose to become under-employed in order to get better jobs later. However, the permanence of under-

[2]Since we take the education level as given and study the resulting worker allocation problem, our focus is different from the over-education literature which mainly studies the returns from schooling.

[3]See Verdugo and Verdugo (1989); Blázquez and Jansen (2008); Leuven and Oosterbeek (2011); Quintini (2011) for more recent evidence.

employment suggests that another mechanism is also at play, and our modeling of under-employment revives an older idea from Thurow (1975), in which individuals compete against one another for job opportunities, and in which higher educated workers can crowd out lower educated workers.

Our approach to modeling under-employment builds on the search and matching literature with multiple islands and heterogeneous agents,[4] and on the competitive search literature with heterogeneous agents (Eeckhout and Kircher, 2010), in which firms post wage offers and workers can direct their search to the markets with the most attractive alternatives. In contrast to these papers, we relax the assumption of random matching. Our modeling of non-random hiring, in which firms can choose among different applicants, builds on the intuition of Blanchard and Diamond (1994). The models closest to ours are Shi (2001, 2002) and Shimer (2005) in which a vacancy can also simultaneously receive multiple applications from heterogeneous job seekers. In contrast with Shi (2001), we relax the assumption that firms commit ex-ante to hire one particular type of job seekers. In contrast with Shi (2002) and Shimer (2005), we relax the wage-posting assumption that firms commit to a wage and cannot negotiate a lower wage when they receive multiple applications of the same type.

Our modeling of wage negotiation is related to the competing-auction theories of Shimer (1999) and Julien, Kennes and King (2000), in which job candidates auction their labor services to employers. We propose a tractable bargaining setup that can capture wage negotiations with (i) multiple, and (ii) heterogeneous, applicants in a non-random hiring setting. In our setup, 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 of the other applicants.

The counter-cyclicality of under-employment echoes the counter-cyclicality of mismatch (Sahin et al., 2012; Barnichon and Figura, 2013), and our paper relates to recent theoretical work on worker mobility across occupations or industries over the business cycle (Alvarez and Shimer, 2011; Carrillo-Tudela and Visschers, 2013). We differ from that literature in two dimensions. First, we focus on vertical mobility –between high-degree requirements and low-degree requirements occupations–. Second, in our framework, the decision of high-skill workers to move across islands is driven by their ranking advantage, and cyclical fluctuations in mismatch can arise

[4]See e.g., Albrecht and Vroman (2002), Gautier (2002), Dolado, Jansen and Jimeno (2009), Charlot and Decreuse (2010).

without time-varying heterogeneity across islands, i.e., even if labor market tightness is equal across islands.

Finally, the idea that unemployment may trickle-down to lower layers received some empirical support in Gautier et al. (2002) and more recently Beaudry, Green and Sand (2013). In particular, Beaudry, Green and Sand (2013) 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 1, we study the properties of under-employment. In Section 2, we argue that under-employment is a puzzle for standard search models. Sections 3 and 4 present a model of under-employment, first in partial equilibrium and then in general equilibrium. Section 5 discusses the optimality of the decentralized allocation. Section 6 brings the model to the data, and the final section concludes.

1 The anatomy of under-employment

This section presents three stylized facts about under-employment in the US.

First, under-employment is highly counter-cyclical, and fluctuations in under-employment are large, comparable with fluctuations in unemployment. Thus, business cycle fluctuations affect the labor market not only through the unemployment rate but also through the under-employment rate. Second, under-employment can be a persistent state for workers who decide to move down the job ladder. Third, moving down the job ladder implies a substantial wage loss, although high-educated workers earn a premium over low-educated workers employed in the same occupation.

1.1 The cyclicality of under-employment

We define as under-employed an individual with some college (or more) who is employed in an occupation that requires at most a high-school degree.[5]

Occupations are defined at the 3-digit Standard Occupational Classification (SOC) level. To measure a worker's education level and occupation, we use micro-data from the CPS between 1979 and 2013. To measure the degree requirement of an occupation, we use data from the BLS 2012 Occupational Outlook Handbook on the

[5]Changing the education threshold, for instance by defining as under-employed an individual with a college degree working in an occupation requiring "some college but no degree", gives very similar conclusions and stylized facts. Only the average level of under-employment is changed.

education requirements by occupation. The BLS determines the education requirement of each occupation from federal and state regulations and from the typical path of entry into a job. We keep the education requirements by occupation fixed over time.

Figure 1 plots the US under-employment rate –the fraction of working college-educated individuals who are under-employed–, cleared from compositional effects.[6] We can see that under-employment is strongly counter-cyclical, increasing in recessions and periods of slack labor market.[7] For instance, in the last recession the under-employment rate increased by about 3 percentage points between 2009 and 2012, which amounts to about 3 million additional under-employed workers. This number is comparable with the increase in the number of unemployed of about 7 million people over the same time period. In other words, the increase in unemployment during a recession is only one aspect of the effect of a recession on the labor market. Another substantial, but so far overlooked, aspect is the large increase in the fraction of workers who are under-employed, in line with our opening quote from the Wall Street Journal.

1.2 The permanence of under-employment

We now show using CPS micro data that becoming under-employed is not a transitory stage for high-educated workers. Instead, more than 70% of high-educated workers who move down the job ladder are still under-employed (or out-of-job) one year later.

In the CPS, an individual is surveyed for four consecutive months, left out for eight months, and then surveyed again for four consecutive months, allowing us to observe the same individual across 8 surveys. To evaluate the persistence of under-employment, we exploit the panel dimension of the CPS, and we measure the fraction of "newly-under-employed" workers who are still under-employed one year later.

More specifically, we proceed in two steps.

[6]Fluctuations in the under-employment rate could be due to demographics or compositional industry effects. For instance, if certain industries feature more under-employment than others, and if these industries have more cyclical employment, the under-employment rate would appear to be cyclical. To control for such compositional changes, Figure 1 is constructed after controlling for a number of observable characteristics. Specifically, we regress a dummy capturing whether an individual is under-employed on a set of industry fixed effects (defined at the 3-digit NAICS level), seasonal dummies, age and sex of the surveyed individual, state-fixed effects, and we plot the residual of this regression centered at the mean of the under-employment rate over the sample period. Because of changes in industry group definitions in 1983, the under-employment rate reported in figure 1 only starts in 1983.

[7]Note also how under-employment lags the unemployment rate.

In a first step, we use the first four surveyed months to identify "newly- under-employed" individuals. We define as "newly-under-employed" an individual who (i) is unemployed in his first surveyed months, (ii) reports a previous occupation in line with his education level,[8] and (iii) finds an under-employed job during the first four surveyed months.

In a second step, we use the four surveys conducted after the eight-month break to observe the under-employed's employment status exactly one year after becoming under-employed.

As shown in table 1, only 28% of newly-under-employed workers moved back up the job ladder within a year, indicating that under-employment does not appear to be a strong jumping board for better jobs. Instead, under-employment appears to be a persistent state for many workers.[9]

1.3 The wage of under-employment

We now argue that becoming under-employed likely involves a substantial wage loss –an under-employed worker earning about 30 percent less than his non-under-employed counterpart–, and that this "wage cost" is higher in times of high unemployment. Moreover, despite this wage loss, high-educated workers earn a premium over low-educated workers employed in the same occupation.

The wage cost of under-employment We first focus on the wage cost of under-employment. To measure such a wage cost, we need to compare the hiring wage of under-employed individuals with the hiring wage of *identical* individuals who find occupations in line with their education (individuals referred to as "well-employed" hereafter). Unfortunately, this thought experiment is not possible because of unobserved heterogeneity: If individuals with lower (unobserved) ability select themselves into under-employment more often than individuals with higher (unobserved) ability, the hiring wage gap between under-employed and well-employed individuals will also reflect unobserved heterogeneity in ability instead of the true wage cost of under-employment.

In order to get as close as possible to this ideal thought experiment (within the

[8]The CPS asks job seekers about their previous occupation, which, combined with the reported education level, allows us to evaluate whether an individual was previously employed in an occupation matching his education level.

[9]Assuming that labor market transitions follow a Markov process, this fraction implies that the average time to wait before moving up the job ladder for a newly-under-employed worker is 1/0.28, which amounts to about 3.5 years. This finding is in line with independent evidence from Verhaest and Schatteman (2010) who follow school leavers for seven years after their entry into the labor market and find that under-employment is a persistent state.

constraints of the CPS) and evaluate the wage cost of under-employment, we proceed as follows:

We use data from the CPS Merged Outgoing Rotation Groups (MORG) over 1979-2013, and we consider the following baseline equation

$$\omega_{ijt} = \alpha + \beta D^{ue}_{ijt} + \gamma X_{it} + \mu_t + \varepsilon_{ijt} \tag{1}$$

in which we model the hiring wage of a highly educated (some college or more) unemployed individual of type i finding a job of type j in period t.

ω_{ijt}, our dependent variable, is the logarithm of the real hourly wage of new hires.[10] D^{ue}_{ijt} is a dummy equal to 1 if the newly-hired individual is under-employed, and the coefficient β is our variable of interest, meant to capture the "wage cost" of under-employment. The vector X_{it} includes observable individual characteristics, and μ_t is a time fixed effect to control for cyclical variations.

We estimate (1) using only newly-under-employed individuals, so that the wage gap is not driven by the presence of (unobserved) low-ability workers who are always under-employed.[11] Specifically, and similarly to the previous section, we restrict ourselves to *newly*-under-employed individuals, i.e., individuals who (i) are unemployed in a given surveyed month, (ii) report a previous occupation in line with their education level, and (iii) find an under-employed job during the next surveyed month.

We then consider two specifications.

In the the first specification, we evaluate the wage gap after controlling for a range of observables X_{it} meant to capture selection into under-employment. Specifically, we include the usual observable characteristics (age, sex, unemployment duration, state of residence) as well as previous occupation fixed effects. The use of previous-occupation fixed effects implies that we are comparing the hiring wage of individuals who used to work in the same occupation (defined at the 3-digit level).[12] By focusing on narrowly defined occupations, the goal is to restrict the sample to similar individuals.

[10] The hiring wage is the hourly wage deflated by the Personal Consumption Expenditures (PCE) deflator.

[11] Indeed, the group of under-employed workers is likely composed of workers with different unobserved characteristics. In particular, some low unobserved ability (or low college quality) college-educated workers may be "permanently" under-employed, i.e., always employed in jobs that do not require some college. Since these low-ability workers will always receive a lower wage than high-ability workers, looking at the unconditional wage difference between well-employed and under-employed workers would bias our result towards a high wage cost of under-employment. To address this issue, we only focus on "marginal" under-employed workers, that is workers who used to be well-employed and just moved down the job ladder.

[12] There are around 500 occupations at the 3-digit level in the CPS, so that the use of previous occupation fixed effects allows us to compare individuals who used to work in similar occupations.

However, this first specification could still be affected by a selection on unobservables problem, since unemployment duration or previous occupation fixed effects may not control for all of the unobserved heterogeneity across individuals. In a second specification, we thus add an additional control: the wage in the previous occupation.[13] Assuming that wage differences reflect unobserved heterogeneity (given that we already control for the usual observable characteristics), controlling for the previous wage will help us control for part of the unobserved selection into under-employment.

The first panel of table 2 presents the results. As shown in the first column of table 2, the unconditional hiring wage gap between under-employed and well-employed individuals is about 42 percent. Controlling for age, sex, unemployment duration and previous occupation fixed effects reduces the estimate only slightly to 36 percent (second column). After including past wages, the estimate falls some more but still remains large and significant at 32 percent (third column), implying that the wage cost of under-employment is likely to be substantial.

Finally, we study the cyclicality of the wage cost of under-employment. Although we cannot evaluate the cyclicality of the wage cost of under-employment using our preferred specification that controls for the previous wage, we can still use our first specification with all the other controls (age, sex, unemployment duration, state of residence, and previous occupation fixed effects).[14] As shown in table 3, the wage cost of under-employment is significantly higher in times of high unemployment.[15]

Wage premium We now study the "wage premium" of under-employment, i.e., the difference between the wage of an under-employed high-educated worker and the wage of a low-educated worker employed in the *same* occupation, defined at the three-digit level.

We use, as before, data from the CPS MORG over 1979-2013, and we consider

[13]This is however only possible for a subset of the sample. Since wages are only observed in the 4th and 8th survey, we need to (i) match individuals across the eight surveys (as in the previous section on the permanence of under-employment), (ii) isolate individuals who are well-employed (i.e., in an occupation matching their education level) in the 4th survey, and (iii) restrict ourselves to individuals who found a job between the 7th and 8th survey (since we are focusing on the hiring wage).

[14]Because of sample size, we cannot assess the cyclicality of the wage cost/premium of under-employment using only newly-under-employed workers (as used in table 2). We instead use the hiring wage of all under-employed workers.

[15]Although cyclical movements in the wage cost of under-employment could be driven by cyclical variations in the (unobserved) ability of the average job seeker, recent evidence suggests that this is unlikely. Mueller (2012) shows that in recessions the pool of unemployed shifts towards workers with high wages in their previous job. This indicates that the average ability of the pool of applicants for low-tech jobs increases in recessions, which would make the wage cost of under-employment pro-cyclical, not counter-cyclical.

the following specification:

$$\omega_{ikt} = a + bD^c_{it} + cX_{it} + \nu_k + d_t + e_{ikt} \quad (2)$$

where i indexes an unemployed worker finding an occupation k requiring low education in period t, and D^c_{it} is a dummy equal to one if the worker has a college degree. In this specification, we control for invariant occupation-specific characteristics (fixed effects ν_k for three-digit occupational code). The coefficient b then captures the worker-specific wage premium associated with a college education.

As shown in the second panel of table 2, we find that high-educated and low-educated workers hired in the same occupation are treated differently by firms: high-educated workers receive a premium in their hiring wage of about 25 percent compared to low-educated workers.[16]

2 Under-employment in standard search models

The previous section highlighted three new stylized facts about the labor market: (i) a substantial fraction of workers are under-employed and the under-employment rate is counter-cyclical, (ii) under-employment is a persistent state for workers, and (iii) under-employment involves a substantial wage loss but high-educated workers earn a premium over their low-educated colleagues.

In this section, we first argue that the existence and cyclicality of under-employment are difficult to rationalize in modern search-based models of the labor market, whether it is the search and matching model of Pissarides (1985) or a competitive search model à la Moen (1997), because in those models, matching is random and independent of the worker type, in the sense that any two workers searching in the *same* labor market (and with the same search intensity) are equally likely to form a match. We then discuss how departing from random matching can help rationalize the existence and cyclicality of under-employment.

2.1 Under-employment and random matching

Consider a labor market divided in two islands; a high-tech island (island 2) and a low-tech island (island 1). In each island, and this is a key assumption in standard search models, matching is random and the matching probability is independent of the worker type.

[16] This number is consistent with previous findings on the effect of over-education on wages (Duncan and Hoffman, 1981).

Denote f_1 the job finding rate of a high-skill worker in island 1 and U_1 (resp. W_1) the value of being unemployed (resp. employed) in island 1, and f_2 and U_2 (W_2) the corresponding values in the high-tech island 2.

As in the search and matching literature with multiple islands (Albrecht and Vroman, 2002) or in the competitive search literature (Moen, 1997), high-skill workers choose in which labor market to look for a job. If under-employment exists in equilibrium, high skill workers must be indifferent between all islands, which implies $U_1 = U_2$ or:[17]

$$\frac{f_1}{r + \lambda_1 + f_1}(w_1 - b) = \frac{f_2}{r + \lambda_2 + f_2}(w_2 - b) \tag{3}$$

in which b denotes the value of home production including the unemployment benefits, λ_i denotes the job separation rate in island i and r denotes the rate at which the future is discounted.

Since under-employment entails a wage loss for high-skill worker ($w_1 < w_2$), then high-skill workers must have a higher job finding rate in island 1 than in island 2 ($f_1 > f_2$).[18] When matching is random, the job finding rate is identical across worker types, so that $f_1 > f_2$ implies that it is easier to find a job in a low-requirement occupation *regardless* of one's type. However, as illustrated in Figure 2, this prediction is not supported by the data: finding a job in a low-requirement occupation takes, on average, just as long as finding a job in a high-requirement occupation. In other words, in the context of a standard search model, the existence of under-employment cannot be easily reconciled with the fact that the low-tech island has a higher job finding rate than the high tech island.

The higher level of under-employment in recessions (i.e., the counter-cyclicality of under-emploment) is also difficult to rationalize with standard search models, because the cost of becoming under-employed increases in recessions: (i) the wage cost of under-employment increases (3), and (ii) the time to find a job in a low-requirement occupation increases more than the time to find a job in a high-requirement occupation (Figure 2).

2.2 Matching with ranking

To explain the existence and cyclicality of under-employment, this paper proposes to depart from the hypothesis that matching is strictly random. Instead, we propose a

[17]Considering the allocation in steady-state and assuming for ease of exposition that unemployment benefits are identical across islands and that separation rates are identical across islands. Recall that in an island with random search and constant job separation rate λ_i, the value of working in island i is given by $rW_i = w_i + \lambda_i (U_i - W_i)$ and $rU_i = b + f_i (W_i - U_i)$, which gives $rU_i = \frac{f_i}{r+\lambda_i+f_i}(w_i - b)$.

[18]The separation rates are too small to matter ($\lambda_i << f_i$, Shimer (2012)).

search model with an alternative matching process in which firms can rank applicants and hire their preferred candidate.

In addition to being intuitively appealing, we show in the Appendix that the idea of ranking is supported by experimental data compiled by Kroft, Lange and Notowidigdo (2013). Using fictitious job applications in which both the quality of individual applications and the quality of the application pool are randomized, we find that the rank of a candidate in a queue of applications influences his call-back rate beyond the sheer quality of his resume, consistent with ranking.

The rest of the paper will develop a theoretical framework where ranking arises endogenously as the outcome of a wage bargaining process with multiple job applicants. Before however, we provide some intuition for why ranking can help rationalize (i) the existence and (ii) the counter-cyclicality of equilibrium under-unemployment.

First, with ranking, even if the islands are equally congested, high-educated workers have a higher job finding rate in the low-tech island than in the high-tech island. To see it, notice that the ranking advantage gives high-skill workers an ability to partially jump the unemployment queue and move in front of lower-skilled applicants (since they are hired over any competing lower-skilled job applicants). However, by how much they can jump the queue depends on the share of low-skill workers in each island. If there are more low-skill job seekers in the low-tech island (as will typically be the case in the model), high-skill workers have a stronger ability to jump the queue in the low-tech island than in the high-tech island, which means that they can find a job faster by becoming under-employed. As a result, the arbitrage equation (3) can be verified and under-employment can exist in equilibrium, because the wage loss associated with under-employment is compensated by a shorter unemployment spell.

Second, ranking can help rationalize why under-employment increases during recessions. Imagine that an adverse aggregate shock lowers vacancy posting in both islands, so that the initial unemployment queues –the initial job seekers to job openings ratios– in each island increase by the same amount. If high-skill workers have a stronger ability to jump the queue in the low-tech island than in the high-tech island, the size of the queue will have a smaller influence on their job finding rate in the low-tech island. As a result, high-skill workers will be less affected by the longer unemployment queue in the low-tech island,[19] and they can smooth the adverse aggregate shock by moving down the job ladder in greater proportion.

[19] While this statement is true and helpful for intuition, it does not take into account the change in the expected wage in the low-tech island and how that change also influences high-skills' decision to move down the ladder. The model will allow us to spell out these different channels.

3 A model of under-employment

This section presents a model of under-employment and describes its three key ingredients: (i) heterogeneity across workers and jobs, (ii) coordination frictions, and (iii) wage competition between workers.

3.1 Preferences, technology and market structure

For clarity of exposition, the model is static and consists of one period. However, in the quantitative section, we generalize the model to a dynamic and stochastic setting.

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 T_n. Islands are indexed such that island N has the highest technology level, and island 1 the lowest. 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.[20] Workers of type N are the most productive. 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.[21]

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.

A firm with technology T_i paired with a worker of type j produces $\varphi_{i,j}$ for any $(i, j) \in \{1, \dots, N\}^2$. Unmatched jobs and workers produce nothing.

For ease of presentation, in the paper we will only focus on the case with $N = 2$ –two islands (i.e., two firm types) and two worker types–, since the $N = 2$ case already contains the most interesting lessons from our model. Characterizing the equilibrium in a model with $N > 2$ is relatively straightforward and is discussed in a separate online Appendix.

[20]When going from the measurement of under-employment to a theory of under-employment, we need to take a stand on the returns to education. We posit that education provides general skills that confer higher-educated workers a productive advantage over lower-educated workers in any job. Doing so, we ignore the role of specific skills in certain occupations.

[21]As shown in the Appendix for the dynamic model, the assumption of no unemployment benefits/home production can be easily relaxed and is only used here for analytical simplicity.

Finally, in order to ensure assortative matching in equilibrium, we need a strong degree of complementarity between workers' skill and firms' technology –a form of super-modularity in this context, i.e., $\frac{\varphi_{1,2}}{\varphi_{1,1}} < \frac{\varphi_{2,2}}{\varphi_{2,1}}$.

3.2 Coordination frictions

Each worker must settle to one island and can apply to at most one job.[22] 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.

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), in which each application (ball) is randomly and independently allocated to a vacancy (urn). With a large number V of vacancies and numbers L_1 and L_2 of type-1 and type-2 applicants, the probability $P(a_1, a_2)$ that a firm faces a_1 applicants of types 1 and a_2 of type 2 follows a multinomial distribution which can be approximated with a Poisson distribution

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

3.3 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. The outcome of our wage bargaining collapses to the standard Nash bargaining outcome when the firm negotiates with only one applicant.

There is perfect information, and all agents observe the pool of applicants and their types. With probability $1 - \beta$, the firm makes a take-it-or-leave-it offer to its preferred candidate, and with probability β each candidate makes a take-it-or-leave-it offer to the firm. When the firm faces only one candidate, the expected payoffs are similar to a standard Nash bargaining. With more than one applicant, however,

[22]This assumption implies that the worker's choice variable is which island to send an application to. An observationally equivalent formulation of our model would be to assume that each worker can simultaneously search across different islands but must choose how to allocate his (finite) search time across islands. For instance, each worker could send one application to an island at random, but with the probability of sending to a given island being the choice variable.

candidates may compete against each other and bid down their wage.[23]

When the firm makes the take-it-or-leave-it offer, it captures all the surplus: the best candidate is hired with a wage equal to the worker's outside option. If there are more than one applicant of the best type, the firm picks one at random. When job applicants make the take-it-or-leave-it offer, the best applicant makes an offer in which the firm captures all the surplus above what would be produced by the second-best candidate.

Denote φ_2 the output generated by the highest-type applicant and $\varphi_1 \leq \varphi_2$ the output generated by the second-best applicant (with the convention that $\varphi_1 = 0$ if there is no second applicant). The outcome of the wage bargaining process is then

$$\omega = \beta(\varphi_2 - \varphi_1).$$

When the two best candidates are identical and $\varphi_1 = \varphi_2$, as would be the case with homogeneous applicants, the firm gets all the surplus from the match, and the hired candidate (chosen at random among the applicants) gets zero.

When there is only one applicant and $\varphi_1 = 0$, the applicant receives $\omega = \beta\varphi_2$, i.e., a share β of the surplus as in the standard Nash bargaining case.

3.4 Timing

The timing of events is as follows. (i) 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; (ii) In each island, applications are randomly allocated to vacancies; (iii) A wage negotiation ensues between the firm and its (possibly multiple) applicants; (iv) Firm-worker matches are formed and production starts. Finally, firms pay workers and realize profits.

4 Equilibrium

In this section, we describe the equilibrium allocation characterizing our economy, and we present some comparative statics exercises to illustrate how under-

[23]Our assumptions provide a very tractable wage bargaining rule. Allowing for a more general negotiation wage rule would complicate the analysis but would not affect the 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 identical: the best-applicant would get the job and his wage would depend negatively on the productivity and number of other applicants.

employment behaves in response to fluctuations in labor demand.

4.1 Partial equilibrium with exogenous labor demand

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

Definition 1. *Partial Equilibrium Allocation.*

Workers make optimal decisions. Each worker of each type $i \in \{1,2\}$ decides on which island $j \in \{1,2\}$ to search for a job. The equilibrium is given by a mapping of such worker choices $C_i : [0, L_i] \mapsto \{1,2\}$ for each type i.

Denote x_2 the share of type 2 workers searching in island 1, denote $E\omega_{n,m}$ the expected income of an individual of type n searching in island m, and denote $h_2 = \frac{L_2}{L_1}$ the size of the pool of type 2 workers relative to that of type 1. For instance, a low h_2 captures a situation in which the distribution of worker types is pyramidal: there are few high types and many low types.

The following proposition characterizes the equilibrium allocation in partial equilibrium:[24]

Proposition 1. *There is a unique equilibrium allocation of workers satisfying*

- *Type 2 workers are indifferent between islands 1 and 2, and x_2 is given by the arbitrage condition*

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

with

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

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

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

Proof. Online Appendix. □

[24] We only focus on economies with a positive rate of under-employment in equilibrium (which is always unique). The online Appendix describes the conditions that ensure the existence of a strictly positive rate of under-employment. Intuitively, we need to impose that the number of vacancies in island 2 is not too large, i.e., $E\omega_{2,2} < E\omega_{2,1}$ when $x_2 = 0$. Alternatively, in a model with a strictly positive rate of over-employment and no under-employment, our endogenous ranking mechanism would occur in the high-tech island. The main difference implied by this different configuration is that unemployment would then trickle up.

Each worker searches for a job in the island that provides him with the highest expected wage, and in equilibrium, a high-skill (type 2) worker is indifferent between looking for a job in the high-tech island (island 2) and looking for a job in the low-tech island (island 1), while a low-skill (type 1) worker strictly prefers looking for a job in the low-tech island (island 1). The arbitrage condition, $A(x_2) = 0$, determines, x_2, the equilibrium allocation of type 2 workers across the two islands.

Figure 3 depicts the equilibrium allocation of high-skill 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 high-skill workers searching in island 1 lowers congestion in island 2, which lessens the competition high-skill workers face in island 2 and increases the expected wage. In contrast, an increase in the fraction of high-skill workers searching in island 1 makes island 1 more congested, which increases the competition workers face in island 1 and lowers the expected wage. Intuitively, the expected income of high skill workers is driven by their uniqueness, as it determines both their ability to find a job easily (by being preferably hired over low skill workers) and their ability to obtain a wage premium over low skill workers. As the number of high skill workers increase, each high-skill worker becomes less unique and thus has a lower job finding rate (facing more competition from their peers) and receives a lower wage premium.

4.2 Implications of under-employment

In this section, we present two comparative statics exercises that highlight two interesting properties of our model of under-employment.

First, we illustrate how under-employment can have (regressive) distributional implications. When high-skill workers move down the job ladder, they take the jobs of low-skill individuals. Through this process, unemployment trickles down from the upper-occupation groups to the lower-occupation groups.

Second, we show how aggregate shocks generate counter-cyclical movements in under-employment.

The trickle-down of unemployment Consider an adverse labor demand shock affecting only the high-tech island, i.e., an increase in the queue length q_2. As shown in figure 4, it shifts down the the $E\omega_{2,2}$ curve –the expected wage earned in island 2–, and generates a higher equilibrium x_2 as high-skill workers use under-employment to smooth the shock and the associated decline in expected income. As a result, low-skill workers see a decrease in their expected earnings (as shown by the $E\omega_{1,1}$ curve), because they face more competition from high-skill workers searching in island 1. In other words, a shock affecting the high type group *trickles down* to the lower type group.

The counter-cyclicality of unemployment Consider now an adverse aggregate shock that affects both islands equally. Specifically, the shock is such that absent any change in the under-employment rate of high-skill workers, both worker types would be equally affected by the shock. Then, one can show that under-employment unambiguously increases. The following corollary formalizes this result:

Corollary 1. *Consider an adverse aggregate shock $\Delta q > 0$ that affects the queue lengths in each island such that*

$$\Delta q_2(1 - x_2) = \Delta q_1(h_2 x_2 + 1) > 0$$

Then, the level of under-employment x_2 increases with

$$\Delta x_2 = \frac{\varphi_{2,1} - \varphi_{1,1}}{\varphi_{2,2}(q_1 h_2 + q_2)} e^{-q_1 x_2 h_2 + q_2(1-x_2)} \Delta q_1 > 0.$$

Proof. Online Appendix. □

To help understand this result, figure 5 plots the corresponding experiment. As can be seen in figure 5, while the aggregate shock shifts the wage curves $E[\omega_{2,2}]$ and $E[\omega_{1,1}]$ by the same amount (by construction), Corollary 1 states that the shift in the $E[\omega_{2,1}]$ curve is less pronounced than the shift in the $E[\omega_{2,2}]$ curve. This reflects the fact that, because of the ranking advantage enjoyed by the high-skills, high-skill workers are less affected by the decrease in job openings in the low-tech island than by the decrease in job openings in the high-tech island. As a result, high-skill workers respond to the adverse aggregate shock by moving down the low-tech island in greater proportion, and under-employment unambiguously increases, i.e., under-employment is counter-cyclical.

Intuitively, our result comes from the fact that high-skill workers can partially jump the queue of unemployed in the low-tech island but *not* in the high-tech island.[25] As a result, the number of job seekers relative to job openings, i.e., the length of the unemployment queue, matters less in the low-tech island, and high-skill workers are less affected by an increase in the queue of unemployed in the

[25] Indeed, because the high-tech island has no low-skill job seekers in equilibrium, high-skill workers cannot benefit from their ranking advantage in the high-tech island.

low-tech island than by the same increase in the high-tech island. High-skill workers can thus smooth an adverse aggregate shock by moving down the job ladder in greater proportion, making under-employment counter-cyclical.

In turn, the counter-cyclicality of under-employment has (regressive) distributional implications. As high-skill workers increase their under-employment rate following the adverse aggregate shock, the wage loss incurred by low-skill workers gets amplified. In other words, while high-skill workers manage to smooth the aggregate labor demand shock, low-skill workers suffer twice, (i) directly from the effect of lower labor demand in island 1 and (ii) indirectly from the increased congestion brought on by the trickle-down of unemployment.

4.3 General equilibrium with Endogenous Labor Demand

In this section, we characterize the general equilibrium (GE) with endogenous labor demand. We leave most of the details for the online Appendix, because the mechanisms isolated in the partial equilibrium analysis carry through to the general equilibrium.

There is an arbitrarily large mass of potential entrants who can settle in island n by adopting technology T_n. A firm still 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.

Definition 2. *General equilibrium allocation with endogenous firm entry.*

Workers and firms make optimal decisions. Each worker of each type $i \in \{1, 2\}$ decides on which island $j \in \{1, 2\}$ to search for a job, and each potential firm entrant in island $j \in \{1, 2\}$ decides whether or not to post a vacancy. The equilibrium is given by worker choices $C_i : [0, L_i] \mapsto \{1, 2\}$ for each worker type i and firm choices $F_j : [0, \infty) \mapsto \{0, 1\}$ for each island j.

The following proposition characterizes the equilibrium allocation in general equilibrium:

Proposition 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.

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

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

Proof. Online Appendix. □

Contrasting the PE and GE allocations Embedding the model in a GE context teaches us two things. First, the basic intuition drawn from the PE case is not over-turned by GE forces. Second, our search model with ranking reduces to a standard search and matching model (Pissarides, 1985) as worker heterogeneity vanishes.

To depict the GE allocation graphically, figure 6 proceeds similarly to figure 3 in the PE case and depicts the equilibrium under-employment rate at 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.[26]

Contrasting figure 6 and figure 3 reveals two differences.

First, the $E\omega_{2,2}$ curve, the expected wage earned in island 2, is now flat and no longer upward slopping as in PE. This result illustrates how "matching with ranking" reduces to random matching (as in Pissarides (1985)) when workers are homogeneous. Indeed, in our model the high-tech island is homogeneous in equilibrium and only populated by high-skill job seekers. In Pissarides (1985), as in our model, the supply of (homogeneous) labor has no effect on the equilibrium queue length: An increase in the number of job seekers raises firms' matching probability, i.e., reduces hiring costs, which leads more firms to enter the market, so that profit and thus the queue length are ultimately unchanged. In other words, job creation always compensates any change in the number of job seekers in a homogeneous island.

Second, the $E\omega_{2,1}$ curve is decreasing in x_2, as in the PE case, but the curve is now less steep. This occurs because of a "quality effect", in that firms respond to changes in the average productivity of the unemployment pool. Specifically, as more high-skill workers search in island 1, this raises firms' probability to meet high-skill applicants (who generate a higher surplus than low skill applicants), which raises firms' profits, and leads to more job creation.[27] As under-employment increases, the

[26] The proofs underlying figure 6 are in the online Appendix.

[27] Such a "quality effect" is well known in random search models with heterogeneous workers.

average productivity of the unemployment pool increases, which fosters firm entry and limits the increase in congestion generated by the inflow of workers. As a result, the wage schedule is flatter with endogenous firm entry.

The quality effect of under-employment With the slight differences between PE and GE in mind, it is easy to see that the redistributive implications of under-employment carry on from the PE case to the GE case.

As in the PE case, an adverse labor demand shock affecting island 2 (a downward shift in the $E\omega_{2,2}$ curve) will trickle down to island 1, as high-skill workers respond to higher congestion in the high-tech island by increasing their under-employment rate, which leads to higher congestion for low-skill workers. The only (qualitative) difference with respect to the PE case, is that, thanks to the increased number of high skill workers in the low tech island, job creation is higher in the low-tech island (the quality effect), and this GE effect dampens the crowding-out coming from the larger number of high-skill workers searching in the low-tech island.

5 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) subject to firms' free entry condition and subject to hiring frictions in each island.[28]

Proposition 3. *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*

$$\begin{aligned} A(x_2^*, q_1^*) &= -E\omega_{2,2} + E\omega_{2,1} \\ &= (1-\beta)\, h_2 q_1^* \varphi_{1,1} e^{-q_1^*(2x_2^* h_2+1)} (\varphi_{2,1} - \varphi_{1,1}) \frac{1}{\frac{\partial \pi_1(x_2^*, q_1^*)}{\partial q_1}} \\ &\geq 0 \end{aligned} \tag{4}$$

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

See e.g., Albrecht and Vroman (2002). However, in a model with random matching, the quality effect leads to an upward slopping wage schedule, so that low skill workers always benefit from under-employment. In our case, the "quality effect" does not lead to an upward slopping wage schedule but makes the wage schedule less downward slopping.

[28]Note that our exercise differs from the usual approach in the search literature (Hosios, 1990), since the planner cannot allocate vacancies across islands. The usual trading externalities are also present in our framework, but we abstract from them in order to focus on the worker allocation problem. An extension of our efficiency discussion would be to consider the problem of a planner allocating *both* workers and vacancies across islands.

If $\beta < 1$ and $\varphi_{2,1} - \varphi_{1,1} > 0$, the decentralized allocation (x_2, q_1) is inefficient and has too much under-employment: $x_2 > x_2^*$.

Proof. Online Appendix. □

The proposition states that the decentralized allocation is, in general, not efficient, and that there is too much under-employment. While we leave a more detailed discussion of efficiency for the online Appendix, we now discuss the intuition behind this result.

As a preliminary step, it is helpful to note that, in this worker allocation problem, maximizing total output while satisfying firms' zero profit condition is the same as maximizing total labor income.[29] Thus, when discussing externalities, we can focus on the effect of a worker's decisions on the expected wage on other job seekers.

When a high-skill worker decides between searching in the high-tech and in the low-tech island, two externalities are present: a congestion externality and a job creation externality.

Consider first the effects of these two externalities in the high-tech island. An increase in the share of high-skill workers in the low-tech island lowers competition between workers in the high-tech island, which ceteris paribus raises the expected wage of job seekers. At the same time, the lower number of job seekers in the high-tech island increases hiring costs for firms, which leads to less job creation and ceteris paribus lowers the expected wage of job seekers. In the high-tech island, workers are homogeneous, and the two externalities exactly compensate each other. As we saw in Section 4, in the high-tech island, job creation always compensates movements in the number of job seekers, and the expected income of high-skill workers is independent of the under-employment rate. Thus, a high-skill worker exerts no externality on the high-tech island.

Consider now the low-tech island, the congestion externality now lowers the expected wages of job seekers (both high-skill and low-skill), while the job creation externality raises expected wages. However, in contrast to what happens in the high-tech island, the two externalities do not compensate each other. Instead, Proposition 3 states that the congestion externality is always stronger than the job creation externality, and firms do not post enough vacancies to compensate for the increased congestion generated by high-skill workers.

With worker heterogeneity, job creation can never fully compensate the congestion externality. While firms' profits increase with the number of high-skill workers,

[29]Indeed, with free entry, we have profits $\pi = y - \omega = c$ so that maximizing total output y is identical to maximizing the wage bill ω.

the presence of low-skill workers limits the increase in profits, because firms face a non-zero probability of receiving applications from only low-skill applicants and then ending up with a low-skill worker and low profit. As a result, job creation does not increase as fast as congestion, and the congestion externality always dominates. It is only in the limiting case when the share of high-skills relative to low-skills goes to one (i.e., when $h_2 x_2 >> 1$ so that high-skill workers are infinitely more numerous than low-skill workers), that the unemployment pool becomes homogeneous again and that the job creation externality exactly compensates the congestion externality.

6 Quantitative exercise

In this section, we build a dynamic stochastic version of our model with under-employment, and we study the properties of the calibrated model. In particular, we evaluate how under-employment can redistribute an aggregate shock from the high-skill groups to the low-skill groups, and how under-employment and job creation interact to shape the labor market experience of low-skill workers. We also explore to what extent aggregate productivity shocks can generate counter-cyclical movements in under-employment, and to what extent under-employment has helped the high skills smooth business cycle fluctuations at the expense of the low skills.

6.1 A dynamic stochastic model with under-employment

To study the quantitative implications of our model, we extend our general equilibrium model of under-employment to a (i) dynamic and (ii) stochastic setting. First, matched workers separate from their job at an exogenous rate λ. Second, productivity is stochastic and subject to shocks, and (log) $\varphi_{1,1,t}$, $\varphi_{2,1,t}$ and $\varphi_{2,2,t}$ follow AR(1) processes with mean (log) $\varphi_{1,1}$, $\varphi_{2,1}$ and $\varphi_{2,2}$. The model is general equilibrium in that a vacancy can freely be created in an island subject to a vacancy posting cost, and this cost is paid in each period.

While we leave a detailed description of this model for the appendix, the main difference with the static model is that workers' reservation wage –the wage making workers indifferent between working at that wage or remaining unemployed– is no longer 0 (or equal to the exogenous value of unemployment benefits). Instead, the reservation wage is endogenous and depends on the labor market allocation, and specifically on workers' distribution across islands and on the job openings to job seekers ratio in each island. The algebra is simplified by the fact that job openings is a jump variable (as in most search models used for business cycle analysis, e.g., Shimer (2005), Gertler and Trigari, 2009), and one can show that the reservation wage of a type i worker is implicitly defined by a familiar relation of the form

$$b_{i,t} = b + \frac{1}{r+\lambda} \int_{b_{i,t}}^{\infty} (w - b_{i,t})\, dF_{i,t}(w) \tag{5}$$

where r denotes the workers' discount rate, b the flow payment of unemployment benefits (or home production) and where F is the (endogenous) distribution of negotiated wages that depends on the job openings to job seekers ratio as well as on the composition of the pool of job seekers.

We simulate the model at a monthly frequency, and Table 4 lists the values of the calibrated parameters. We set the income replacement rate to 40% (Shimer, 2005), workers' bargaining weight to 0.5, and workers' separation rates to match CPS worker-flow data. As detailed in the online Appendix, to use the urn-ball matching function in a quantitative context, we proceed as Blanchard and Diamond (1994) and introduce a matching efficiency term. We set matching efficiency in each island to match the average job finding rates of high-skill and low-skill workers. Normalizing the mean level of $\varphi_{2,2}$ to one, we pin down the mean levels of $\varphi_{1,1}$ and $\varphi_{2,1}$ from the average measured under-employment rate and from the average wage premium earned by over-educated workers ($\frac{\omega_{2,1}}{\omega_{1,1}}$).[30] Finally, the vacancy posting costs are set to match the average job openings to job seekers ratio in typical college-level occupations (professional, services) and in typical high-school level occupations (construction and sales) as measured from the Help Wanted OnLine (HWOL) dataset from the Conference Board. More details are provided in the Appendix.

6.2 Impulse response functions

With our calibrated model, we can study the properties of the model and in particular analyze how under-employment can redistribute aggregate shocks from the high-skill groups to the low-skill groups.

We consider a purely transitory aggregate productivity shock that affects all

[30] Based on this calibration, the model generates a wage loss associated with under-employment of about 10 percent, lower than our estimates from Table 2. The discrepancy can be due to a number of reasons: First, in our empirical exercise, some unobserved heterogeneity likely remains and contributes to inflate our estimate of the wage cost of under-employment. Second, workers who decide to move down the job ladder may under-estimate the wage cost of under-employment or under-estimate how persistent that under-employment can be. Relatedly, workers may not fully appreciate the extent of human capital depreciation associated with under-employment. Third, the discrepancy may also reflect the additional role of on-the-job search. With on-the job-search and human capital depreciation during unemployment, workers who move down the job ladder will accept a larger wage cut than in our model, since being employed preserves from human capital depreciation.

islands equally, i.e., such that $\varphi_{2,2}$, $\varphi_{2,1}$ and $\varphi_{1,1}$ are all lowered by one percent at the beginning of the period $t = 1$. To make the exercise more illustrative, we assume (only for this exercise) that firms post vacancies at the beginning of the period but only observe the composition of the pool of job seekers in the island at the end of the period. As a result, firms only react with a one month lag to the composition of the pool of job seekers in the island. This timing assumption allows us to more clearly separate the (partial equilibrium) crowding-out effect due to the presence of high-skill workers in the low-tech island from the (general equilibrium) quality effect due to the fact that more under-employment leads low-tech firms to post more vacancies.

Figure 7 presents the impulse responses of the variables of interest. At the time of the negative aggregate productivity shock ($t = 1$), job openings decline equally in both islands, and high-skill workers move to the low-tech island to smooth the effect of the shock and under-employment increases (as predicted by Corollary 1). In doing so, they crowd-out the employment opportunities of the low skills, who experience a larger decline in their present discounted value of unemployment (labeled rU, bottom-left panel). Finally, as high-skill workers move down the job ladder, they become less unique, and the wage loss associated with under-employment increases (bottom-right panel).

At $t > 1$, the general equilibrium effect kicks in, and the higher under-employment rate leads low-tech firms to post more vacancies (middle-left panel), which increases the welfare of low-skill job seekers once productivity is back to normal (bottom-left panel).

Finally, Figure 7 highlights an implication of under-employment that could not be seen in the static model. With only a small amount of inertia in firms' job openings decision, purely transitory shocks can have long lasting effects on the economy. Indeed, even though productivity is back to its long-run level after one period, under-employment remains above its long-run level for a long time. This propagation mechanism is due to the general equilibrium effect: the larger number of high-skill workers in the low-tech island (and correspondingly smaller number in the high-tech island) stimulates vacancy posting in the low-tech island and depresses job openings in the low-tech island. This further encourages high-skill workers to remain in the low-tech island even though productivity is back to its steady-state level.

6.3 Simulation

In this section, we explore to what extent aggregate productivity shocks can generate counter-cyclical movements in under-employment and to what extent aggregate shocks are redistributed from the high skills to the low skills.

To get a plausible source of aggregate fluctuations, we simulate an economy with aggregate productivity shocks, where the shocks and persistence of the AR(1) process for $\varphi_{2,2}$, $\varphi_{2,1}$ or $\varphi_{1,1}$ were chosen so that the fluctuations in aggregate output match the fluctuations in Congressional Budget Office (CBO)'s estimate of the output gap.[31]

Figure 8 plots the simulated fluctuations in aggregate productivity (top panel), the implied movements in unemployment and under-employment (second panel), in the value of unemployment (rU, third panel), and in job openings (bottom panel).

We can see from Figure 8 that a model with aggregate productivity shocks can generate counter-cyclical fluctuations in under-employment that are broadly in line with the US experience.

The figure also illustrates how crowding out and job creation interact to shape the labor market experiences of high-skill and low-skill workers. As aggregate productivity fluctuates, high-skill workers smooth the effects of shocks by adjusting the under-employment rate, which crowds out the job opportunities of low-skill job seekers in bad times. As a result, in the two large recessions of the early 1980s and late 2000s, the welfare of low-skill job seekers decreased by 30 percent more than the welfare of high-skill job seekers, and overall the value of unemployment for low-skill workers is about 22 percent *more* volatile than the value of unemployment for high-skill workers (as measured from standard-deviations). All in all, the trickle down of unemployment redistributes aggregate shocks from the high-tech island to the low-tech island, and as a result, high-skill job seekers experience not only a higher value of unemployment ($U_2 > U_1$) but also a substantially lower variance ($sd(U_2) < sd(U_1)$).

Interestingly, this redistributive effect of under-employment is partially offset by firms' reaction to the quality of the pool of job seekers; the quality effect that we discussed in section 4. In recessions, the increase in under-employment raises the average productivity of job seekers in the low-tech island, which leads, ceteris paribus, to more job openings in the low-tech island. As a result, job openings do not decline as much in the low-tech island as in the high-tech island. During booms, the opposite effect is at play: under-employment goes down and job openings in the low-tech island do not increase as much. Through this quality effect, vacancy posting is about 25 percent less volatile in the low tech island than in the high tech

[31] Our model suffers like other search models from the "Shimer puzzle" (Shimer (2005)) and generates too little fluctuations in unemployment given productivity shocks of plausible magnitude. Thus, we do not try to match the volatility of unemployment but concentrate our efforts on simulating an economy with plausible fluctuations in productivity.

island (as measured from standard-deviations), even though both islands suffer the exact same productivity shocks.

Finally, note how the under-employment rate lags the unemployment rate, as in the data (figure 1). The lagging nature of under-employment comes mainly from the fact that the stock of under-employed workers is a slow moving variable, so that the stock of high-skill workers employed in the low tech island changes slowly. Intuitively, the under-employment rate adjusts slowly, because changes in the composition of the pool of employed workers depends on the rate of job separation. Since the probability of employment separation is low, the composition of the unemployment pool changes slowly, and the under-employment rate adjusts slowly to changes in the composition of the unemployment pool. This slow adjustment of the under-employment rate is in contrast with the rapid adjustment of the unemployment rate. The unemployment rate adjusts fast, because the speed of adjustment of unemployment is driven by the job finding rate, which is an order of magnitude larger than the job separation rate.[32]

7 Conclusion

While the unemployment rate is the traditional gauge of the labor market, this paper shows that under-employment is an important, yet overlooked, characteristic of the labor market.

We study empirically the phenomenon of under-employment in the US and show that (i) under-employment is strongly counter-cyclical, (ii) under-employment is costly –an under-employed worker earning about 30 percent less than his non-under-employed counterpart–, and (iii) under-employment can be a persistent state for newly-under-employed individuals.

We then argue that under-employment is hard to rationalize with modern (search-based) models of the labor market, be it a search and matching model à la Pissarides (1985) or a competitive search model à la Moen (1997). We instead propose a search model in which a vacancy can simultaneously receive multiple applications. Through the wage bargaining process, high-skill applicants are systematically hired over less-skilled competing applicants, so that the model generates an endogenous ranking mechanism favoring high-skill workers. In this framework, under-employment (i)

[32]The unemployment rate adjusts quickly to its steady-state level (Shimer, 2012) because the speed of convergence to steady-state is determined mostly by the job finding rate (Barnichon and Nekarda, 2012), which is large in the US. In contrast, one can show that the speed of convergence of the under-employment rate towards its steady-state is dominated by the separation rate, which is an order of magnitude lower than the job finding rate. This explains why the under-employment rate lags the unemployment rate.

exists in equilibrium and (ii) is counter-cyclical.

First, with ranking, under-employment exists, not because low-qualification jobs are more abundant, but because the competition to get a low-qualification job is, from the viewpoint of high-skill workers, less intense. Intuitively, a high-skill worker moves down the occupational ladder, in order to find a job more easily and escape competition from his high-skill peers. Second, with ranking, under-employment is counter-cyclical. Intuitively, following an adverse aggregate labor demand shock affecting all types of jobs, high-skill workers are less affected by the drop in low-skill jobs because of their ranking advantage. As a result, they smooth the aggregate shock by moving down the job ladder in greater proportion, and under-employment increases in recessions.

A quantitative version of the model generates fluctuations in under-employment in line with the US experience and shows that the trickle-down of unemployment can be a non-trivial redistributive mechanism as shocks get redistributed from the high skills to the low skills.

Davis and Wachter (2011) recently showed that the cumulative earnings losses associated with job displacement are (i) substantial and (ii) even larger if the displacement occurs during a recession. While Davis and Wachter (2011) argue that standard search models cannot rationalize such large losses, under-employment may provide an explanation for this finding. Because under-employment is a persistent state, a high-skill worker who moves down the job ladder following displacement will suffer a persistent drop in income. The average cumulative earnings losses upon displacement will be larger during recessions because more high-skill workers will move down the job ladder during recessions. While under-employment is an optimal choice for high-skill workers in our model, it is easy to imagine cases for which high-skill workers are forced down the occupation ladder to a greater extent, because of borrowing constraints or reputation considerations associated with long unemployment spells (Kroft, Lange and Notowidigdo, 2013). Then, if moving back up the ladder is difficult because of asymmetric information or skill deterioration, under-employment opens the door to large welfare costs of business cycle fluctuations.

Finally, the possibility of under-employment opens new, and so far unexplored, returns to 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

Albrecht, J, and S Vroman. 2002. "A Matching Model with Endogenous Skill Requirements." *International Economic Review*, 43(1): 283–305.

Alvarez, Fernando, and Robert Shimer. 2011. "Search and Rest Unemployment." *Econometrica*, 79(1): 75–122.

Barnichon, Regis, and Andrew Figura. 2013. "Labor Market Heterogeneity and the Aggregate Matching Function."

Barron, John M, and John Bishop. 1985. "Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity." *Economic Inquiry*, 23(3): 363–82.

Barron, John M, John Bishop, and William C Dunkelberg. 1985. "Employer Search: The Interviewing and Hiring of New Employees." *The Review of Economics and Statistics*, 67(1): 43–52.

Beaudry, Paul, David A. Green, and Benjamin M. Sand. 2013. "The Great Reversal in the Demand for Skill and Cognitive Tasks." National Bureau of Economic Research, Inc NBER Working Papers 18901.

Behrenz, Lars. 2001. "Who Gets The Job And Why? An Explorative Study Of Employers' Recruitment Behavior." *Journal of Applied Economics*, 0: 255–278.

Blanchard, Olivier Jean, and Peter A Diamond. 1994. "Ranking, Unemployment Duration, and Wages." *Review of Economic Studies*, 61(3): 417–34.

Blázquez, Maite, and Marcel Jansen. 2008. "Search, mismatch and unemployment." *European Economic Review*, 52(3): 498–526.

Butters, G. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *Review of Economic Studies*, 44: 465–491.

Carrillo-Tudela, Carlos, and Ludo Visschers. 2013. "Unemployment and Endogenous Reallocation over the Business Cycle." Institute for the Study of Labor (IZA) IZA Discussion Papers 7124.

Charlot, Olivier, and Bruno Decreuse. 2010. "Over-education for the rich, under-education for the poor: A search-theoretic microfoundation." *Labour Economics*, 17(6): 886–896.

Davis, Steven J., and Till Von Wachter. 2011. "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity*, 43(2 (Fall)): 1–72.

Dolado, Juan, Marcel Jansen, and Juan Jimeno. 2009. "On-the-Job Search in a Matching Model with Heterogeneous Jobs and Workers." *Economic Journal*, 119(534): 200–228.

Duncan, Greg J., and Saul D. Hoffman. 1981. "The incidence and wage effects of overeducation." *Economics of Education Review*, 1(1): 75–86.

Eeckhout, Jan, and Philipp Kircher. 2010. "Sorting and Decentralized Price Competition." *Econometrica*, 78(2): 539–574.

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

Gautier, P. 2002. "Unemployment and Search Externalities in a Model with Heterogeneous Jobs and Workers." *Economica*, 69(273): 21–40.

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

Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies*, 57(2): 279–98.

Julien, Benoit, John Kennes, and Ian King. 2000. "Bidding for Labor." *Review of Economic Dynamics*, 3(4): 619–649.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo. 2013. "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*." *The Quarterly Journal of Economics.*

Leuven, Edwin, and Hessel Oosterbeek. 2011. "Overeducation and Mismatch in the Labor Market." Institute for the Study of Labor (IZA) IZA Discussion Papers 5523.

McGuinness, Seamus. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys*, 20(3): 387–418.

Moen, Espen R. 1997. "Competitive Search Equilibrium." *Journal of Political Economy*, 105(2): 385–411.

Mueller, Andreas I. 2012. "Separations, Sorting and Cyclical Unemployment." Institute for the Study of Labor (IZA) IZA Discussion Papers 6849.

Okun, Arthur M. 1973. "Upward Mobility in a High-Pressure Economy." *Brookings Papers on Economic Activity*, 4(1): 207–262.

Pissarides, Christopher. 1985. "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages." *American Economic Review*, 75(4): 676–90.

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

Quintini, Glenda. 2011. "Right for the Job."

Raza, S. M., and B. N. Carpenter. 1987. "A model of hiring decisions in real employment interviews." *Journal of Applied Psychology*, 72: 596–603.

Reder, M. W. 1955. "The Theory of Occupational Wage Differentials." *The American Economic Review*, 45(5): 833–852.

Sahin, Aysegul, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2012. "Mismatch unemployment." Federal Reserve Bank of New York Staff Reports.

Shimer, Robert. 1999. "Job Auctions."

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

Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics*, 15(2): 127–148.

Sicherman, Nachum. 1991. "Overeducation in the Labor Market." *Journal of Labor Economics*, 9(2): 101–22.

Sicherman, Nachum, and Oded Galor. 1990. "A Theory of Career Mobility." *Journal of Political Economy*, 98(1): 169–92.

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

Verdugo, Richard R., and Naomi Turner Verdugo. 1989. "The Impact of Surplus Schooling on Earnings: Some Additional Findings." *Journal of Human Resources*, 24(4): 629–643.

Verhaest, Dieter, and Tom Schatteman. 2010. "Overeducation in the early career: an analysis using sequence techniques." Working Papers 2010/09.

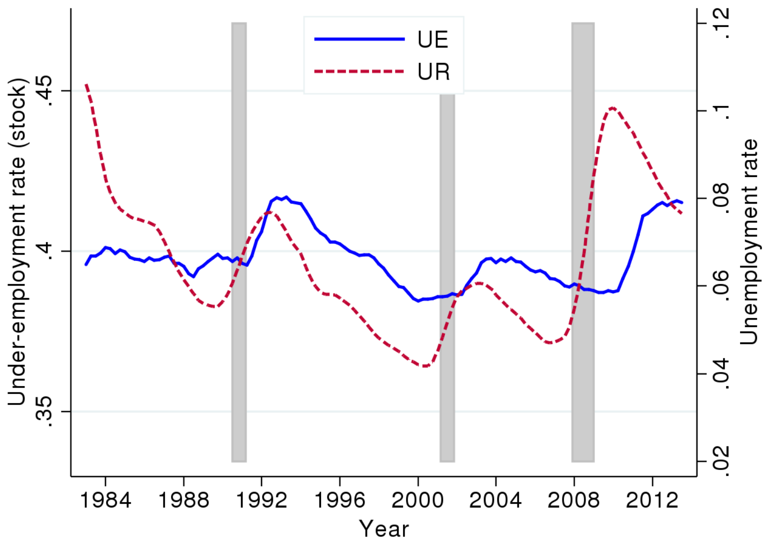

Figure 1. Under-employment in the US

Source: Current Population Survey, 1983-2012. Unemployment rate (dashed line) and Under-employment rate (plain line), defined as the fraction of individuals with some college education (or more) working in occupations requiring at most a high school degree. The under-employment rate series is cleared from compositional effects, as described in the main text.

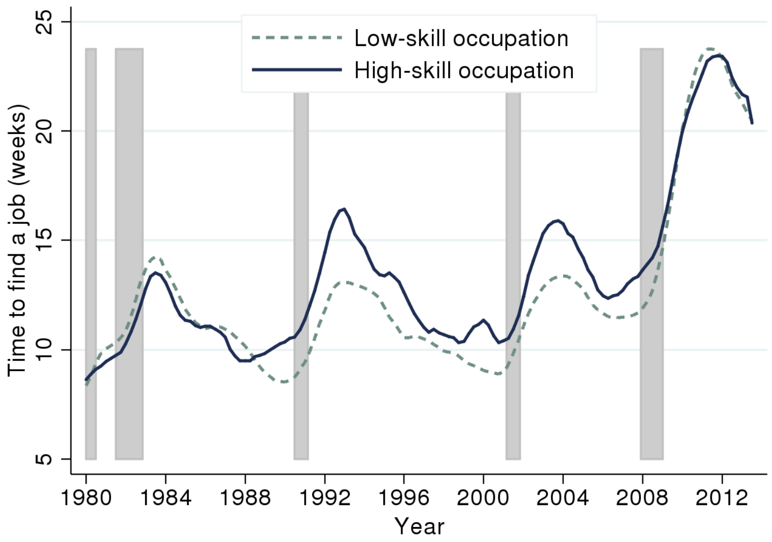

Figure 2. Realized time to find a job (high- versus low-degree requirement occupations).

Source: Current Population Survey, 1980-2012. The realized time to find a job is computed as the average unemployment duration (in weeks) of newly hired workers. The plain line (resp. dashed line) is the realized time to find a job requiring at most a high-school degree (resp. requiring more than a high-school degree).

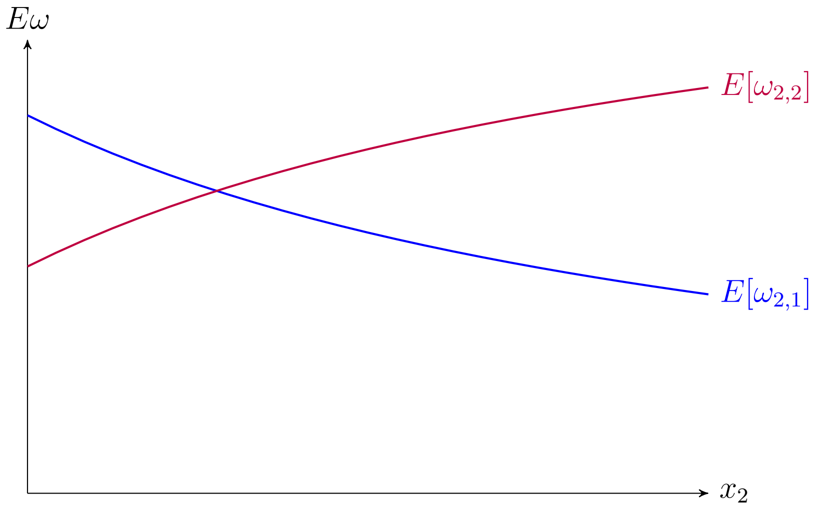

Figure 3. Partial Equilibrium– Expected wage of type-2 workers.

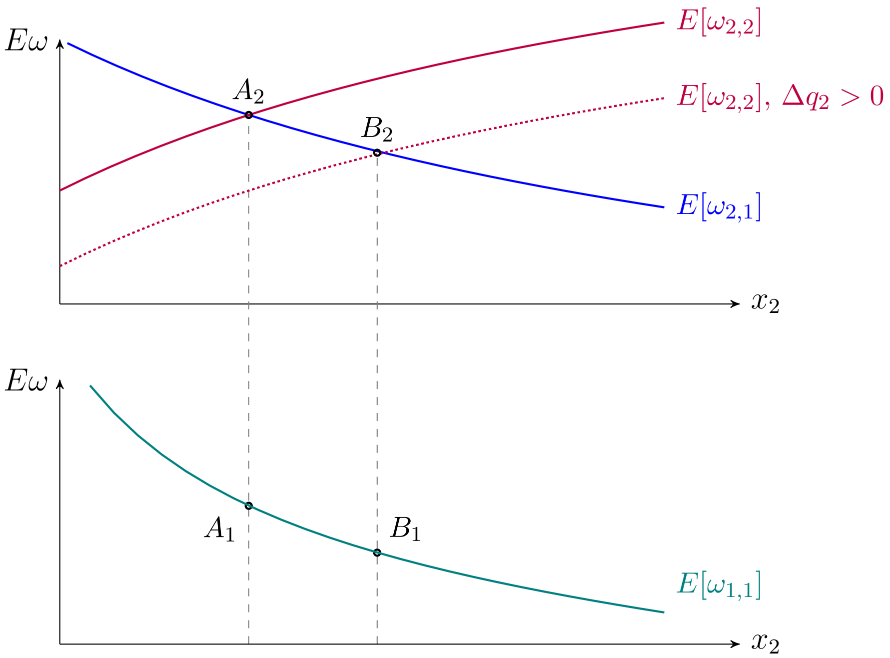

Figure 4. Partial Equilibrium – shock $\Delta q_2 > 0$.

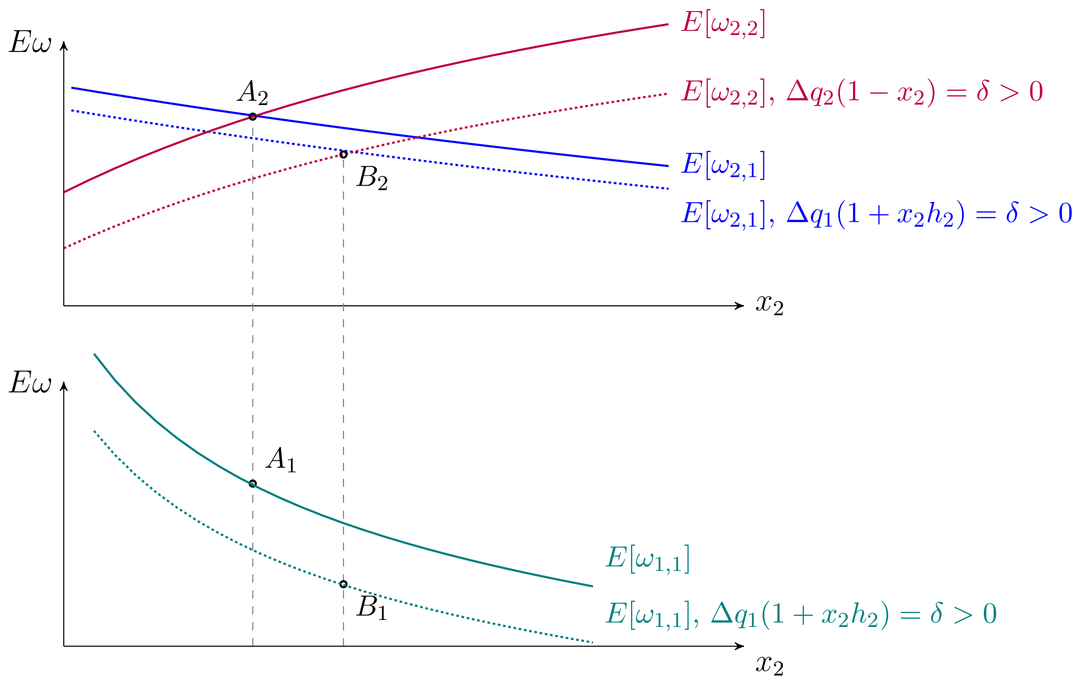

Figure 5. Partial Equilibrium – aggregate shock with $\Delta q_2(1-x_2)=\Delta q_1(1+x_2h_2)>0$.

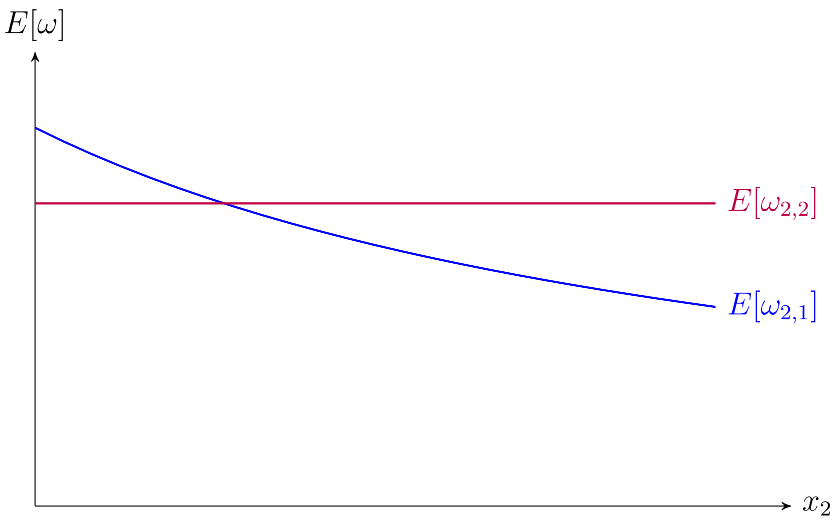

Figure 6. General Equilibrium – Expected wage of type-2 workers.

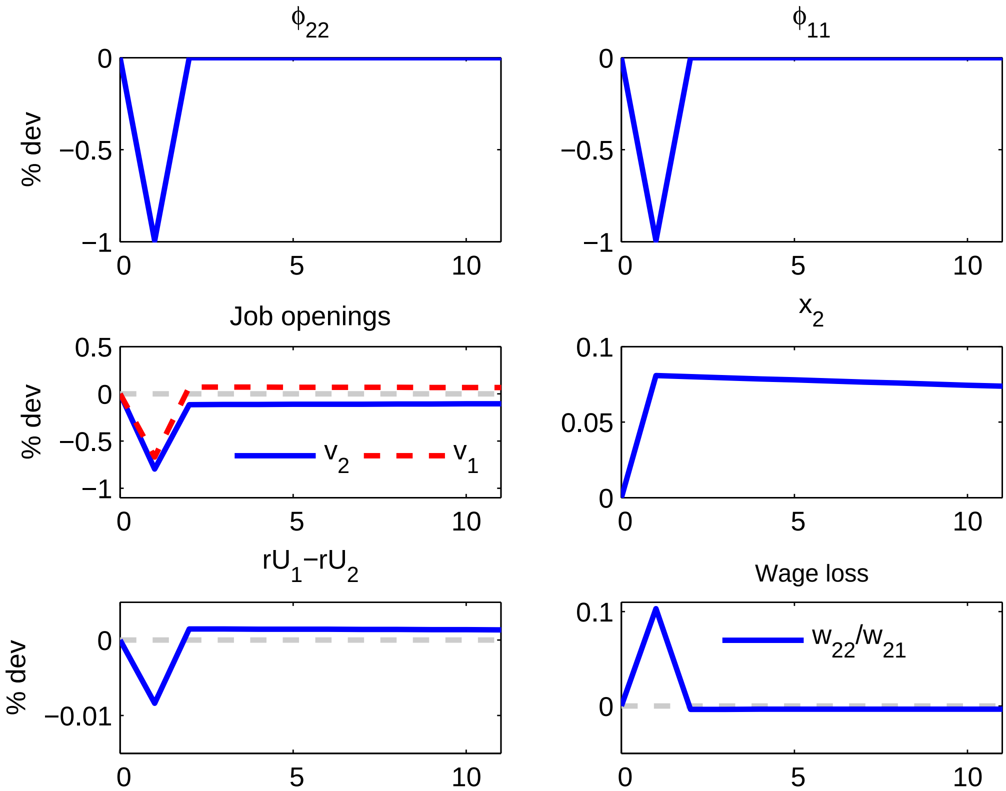

Figure 7. Model-based impulse response functions to an aggregate productivity shock for (i) island-specific productivity ("ϕ", top panels), (ii) job openings (middle-left panel), (iii) fraction of high-skill workers searching in the low-tech island ("x_2", middle-right panel), (iv) difference in the value of unemployment between high-skill and low-skill workers ("$rU_2 - rU_1$", bottom-left panel), and (v) wage cost of under-employment ("w_{22}/w_{21}", bottom-right panel).

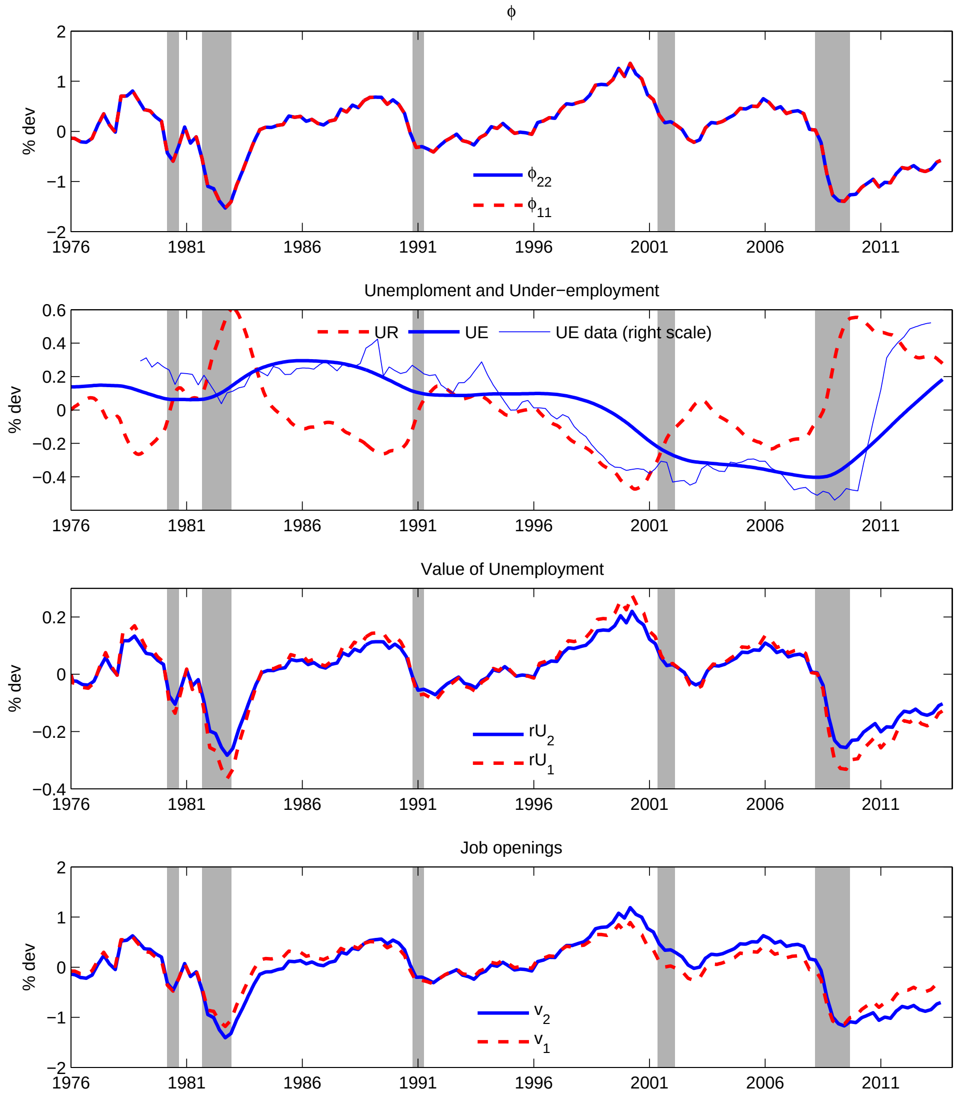

Figure 8. Model simulation. Simulated fluctuations in aggregate productivity (ϕ, top panel), in unemployment and under-employment ("UR", "UE", second panel), in the value of unemployment (rU, third panel), and in job openings ("v", bottom panel).

Table 1. Permanence of under-employment, 1979-2012.

Employment status one year after hiring	E_1	E_2	U
Fraction	.601	.283	.117

Distribution of newly-underemployed workers one year after having been hired in an under-employed occupation. E_1 denotes employment in an occupation requiring at most a high school degree (i.e., under-employment), E_2 denotes employment in an occupation requiring more than a high school degree, and U denotes unemployment. Sample contains 1350 observations.

Table 2. Wage differences by education and occupation degree requirements.

Panel A: Hiring wage of college-educated workers	(1)	(2)	(3)
Under-employed dummy	-.422	-.365	-.317
	(.041)	(.045)	(.042)
State Fixed Effects	Yes	Yes	Yes
Date Fixed Effects	Yes	Yes	Yes
Individual characteristics		Yes	Yes
Past-occupation Fixed Effects		Yes	Yes
Past Wage			Yes
Observations	1,555	1,555	1,555

Panel B: Hiring wage in low degree-requirement occupations	(1)	(2)
College-education dummy	.285	.257
	(.011)	(.011)
State Fixed Effects	Yes	Yes
Date Fixed Effects	Yes	Yes
Individual characteristics		Yes
Current-occupation Fixed Effects	Yes	Yes
Observations	38,636	38,636

Source: CPS (1979-2012). In panel A, the dependent variable is the hiring wage of an unemployed individual with some college, and previously employed in an occupation requiring more than a high-school degree. The under-employment dummy equals 1 if the individual is hired in an occupation requiring a high school degree or less. In panel B, the dependent variable is the hiring wage of an unemployed individual (either newly-under-employed or with low education) in an occupation requiring a high school degree or less. The college-education dummy equals 1 if the individual has some college experience. The controls for individual characteristics are age, sex, state of residence, and duration of unemployment. Robust standard errors are reported in parentheses.

Table 3. Wage differences by education and occupation degree requirements – cyclicality.

Hiring wage of college-educated workers	Full sample (1979-2013)	Recessions (1979-2013)
Under-employed dummy	-.356	-.439
	(.013)	(.065)
State Fixed Effects	Yes	Yes
Date Fixed Effects	Yes	Yes
Individual characteristics	Yes	Yes
Past-occupation Fixed Effects	Yes	Yes
Observations	10,572	995

Source: CPS (1979-2013). The dependent variable is the hiring wage of an unemployed individual with some college. The under-employed dummy equals 1 if the individual is hired in an occupation requiring a high school degree or less. The controls for individual characteristics are age, sex and duration of unemployment. Robust standard errors are reported between parentheses. Recessions are defined as times when the unemployment rate is two standard-deviations above its average level.

Table 4. Parameter values for the benchmark economy.

	Interpretation	Value	Rationale
β	Discount factor	0.9898	Interest rate 4% p.a.
h_2	Dilution	1.10	Unemployment by educ. (CPS 83-13)
β	Bargaining Power	0.5	
$\varphi_{1,1}$	Productivity	0.72	Wages and under-emp. (CPS-MORG 79-13)
$\varphi_{2,1}$	Productivity	0.94	Wages and under-emp. (CPS-MORG 79-13)
$\varphi_{2,2}$	Productivity	1.00	
ν_1	Matching efficiency, low-tech	0.45	Job finding rate (CPS-MORG 79-13)
ν_2	Matching efficiency, high-tech	0.21	Job finding rate (CPS-MORG 79-13)
λ_1	Separation rate, low-tech	0.015	Job separation rate (CPS-MORG 79-13)
λ_2	Separation rate, high-tech	0.01	Job separation rate (CPS-MORG 79-13)
b	Income replacement ratio	0.4	Shimer (2005)

A Appendix: Ranking in the data

In this section, we provide some empirical evidence supporting the notion of ranking and the idea that firms prefer the most qualified workers when facing multiple applicants.[33]

Using experimental data compiled by Kroft, Lange and Notowidigdo (2013), we show that when workers apply for a low-qualification job, workers with stronger resumes are substantially more likely to be called back by the firm than workers with just the required qualifications. Moreover, the rank of the candidate in the queue of applications influences the call-back rate beyond the sheer quality of the resume. This finding is consistent with the idea of ranking, but not with the idea of screening, in which only the quality of the resume should matter.[34]

Kroft, Lange and Notowidigdo (2013) examine the effect of unemployment duration on callback rates using fictitious job applications in which a set of quality indicators are manipulated exogenously. Fictitious resumes were varied in their quality, and four resumes (two of a high type, and two of a low type) were sent to the same job posting. A "low-quality" resume has the minimum qualifications required for the job (in terms of experience and education), while a "high-quality" resume has qualifications that exceeded these minimum, with extra years of experience and an extra level of education.[35]

Using this experimental setup, it is possible to test two important predictions of ranking.

First, over-qualified resumes should face higher call back rates than "just-qualified" resumes competing for the same position. To test this prediction, we estimate three different specifications based on the following equation:

$$C_{i,k} = \alpha + \beta D_i + \delta X_{i,k} + \varepsilon_{i,k}, \tag{6}$$

where i denote an applicant and k denote a job posting. $C_{i,k}$, our dependent variable, is a callback dummy and D_i, our main explaining variable, is a dummy for high-

[33]The notion of ranking is also consistent with qualitative evidence on firms' recruitment behavior, with firms typically interviewing many applicants and hiring the most qualified candidate (Barron, Bishop and Dunkelberg, 1985; Barron and Bishop, 1985; Raza and Carpenter, 1987; Behrenz, 2001).

[34]With screening, hiring is determined by a threshold level above which a candidate is hired (see e.g., models with stochastic job matching, Pissarides, 2000, Chapter 6), and over-qualified candidates are more likely to be above that threshold, irrespective of other candidates. In contrast, with ranking, firms select their new hire among the pool of applicants. Thus, the hiring rate depends on both the number and the quality of the other candidates.

[35]For instance, if the job requires high school completion, the resume would list an associate's degree.

quality resume. The vector of controls $X_{i,k}$ includes individual i characteristics (age and gender) and some job opening k characteristics (type of job and city or job posting fixed effects depending on the specifications).

In the first specification, we only control for the individual characteristics and the three main categories of job (administrative, sales, customer services). In the second specification, we also control for the local labor market by including city fixed effects. Finally, in our third (and preferred) specification, we also include a job posting fixed-effect, which allows us to compare resumes in competition for the same job opening. Table 5 presents the results and shows that, in all 3 specifications, over-qualified resumes have a callback rate that is about 35 percent higher than just-qualified resumes. While the mapping from callback rates to job finding propensities is not straightforward, this evidence strongly suggests that firms do rank candidates by resume's quality, so that higher-educated job seekers likely have an easier time finding a job in a lower requirement occupation than their lower-educated peers.

Second, the rank of an applicant A should matter beyond the sole quality of his resume. In other words, the quality of the other applicants should influence the call-back rate of A above and beyond the quality of A's resume. To exploit variations in the quality of the other applicants, we use a factor that is randomly-drawn across individuals: unemployment duration (see Kroft, Lange and Notowidigdo (2013)). We then test whether the rank (in terms of unemployment duration) of a candidate among the four applicants affect the call-back rate above-and-beyond the resume's quality.[36] The last column of table 5 shows that this is indeed the case, consistent with the implications of ranking. Adding an individual's rank to our baseline specification, we can see that there is an additional, non-negligible, call-back premium for being ranked higher than the other applicants. This result is compatible with the prediction of a model with ranking but not with a model with screening.

[36] We also control for the individual unemployment duration and the average unemployment duration of the three other applicants, as this could also affect the average call-back rate.

Table 5. Callback rate as a function of applicants' type.

Call-back rate	(1)	(2)	(3)	(4)
High-type dummy	.0183***	.0172***	.0176***	.0149**
	(.0056)	(.0055)	(.0060)	(.0062)
Rank				-.0074**
				(.0032)
Individual controls	Yes	Yes	Yes	Yes
Fixed-effects		City	Ad	City
Unemployment duration				Yes
Observations	11,648	11,648	11,648	11,648

Significantly different than zero at * 90% confidence, ** 95% confidence, *** 99% confidence. The dependent variable is a dummy equal to one if the applicant received a call-back. The high-type dummy equals 1 if the resume is of high-quality. Rank denotes how the candidate ranks (in terms of unemployment duration) among the four candidates. The controls are gender, age, individual unemployment duration, average unemployment duration across the four applicants and a set of dummies capturing job type. Standard errors between parentheses are clustered at the ad level. The average callback rates in the sample is .0439.

B Appendix: A dynamic stochastic general equilibrium model with underemployment

In this appendix, we extend our general equilibrium model of under-employment to a (i) dynamic and (ii) stochastic setting. First, in a dynamic setting, matched workers employed in island j separate from their job at an exogenous rate λ_j. Second, the productivity of a match is stochastic and subject to shocks, and productivity in island j for individual i, (log) $\varphi_{ij,t}$, follows an AR(1) process

$$\ln \varphi_{i,j,t} = \rho \ln \varphi_{i,j,t-1} + (1-\rho) \ln \varphi_{i,j} + \varepsilon_t + \varepsilon_{j,t}$$

with $\varphi_{i,j}$ the long-run productivity level, ε_t an aggregate productivity shock affecting all islands equally, and $\varepsilon_{j,t}$ an island specific technology shock.

Third, the model is general equilibrium in that a firm can freely enter an island subject to a vacancy posting cost. The main difference with the static model of the main text is that workers' reservation wage –the wage making workers indifferent between working at that wage or remaining unemployed– is no longer 0 (or the exogenous value of unemployment benefits) in a dynamic setting. Instead, the reservation wage is endogenous and depends on the labor market allocation, specifically workers' distribution across islands and job openings in each island.

B.1 Reservation wage and arbitrage condition

We start by deriving expressions for the endogenous reservation wages. Denote b the flow payment of unemployment benefits (or home production) identical across islands. Let r denote the workers' discount rate. We drop the time indices for the sake of exposure.

The value of being unemployed for a type i looking in island j, U_i, satisfies

$$b_i = rU_i = b + \nu_{i,j} \int \max\left(W_i(w) - U_i, 0\right) dF_i(w)$$

where F is the (endogenous) distributions of negotiated wages.

We now determine the distribution F for type-2 workers, first in island 1 then in island 2.

Conditional on receiving an offer, a type 2 in island 1:

- is the only applicant with probability $e^{-\nu_1 q_1 x_2 h_2} e^{-\nu_1 q_1}$, and receives (in surplus) $\beta \frac{\varphi_{2,1} - b_2}{r+\lambda_1}$.
- is in competition with other type 2 workers with probability $1 - e^{-\nu_1 q_1 x_2 h_2}$, and receives 0 in surplus (he gets exactly b_2).
- is in competition with type 1 workers only and receives in surplus: $\beta(\frac{\varphi_{2,1} - b_2}{r+\lambda_1} - \frac{\varphi_{1,1} - b_1}{r+\lambda_1})$ with probability $e^{-\nu_1 q_1 x_2 h_2} \left(1 - e^{-\nu_1 q_1}\right)$.

We thus get

$$b_2 = rU_2 = b + \nu_1 e^{-\nu_1 q_1 x_2 h_2} e^{-\nu_1 q_1} \beta \frac{\varphi_{2,1} - b_2}{r + \lambda_1}$$
$$+ \nu_1 e^{-\nu_1 q_1 x_2 h_2} \left(1 - e^{-\nu_1 q_1}\right) \beta \left(\frac{\varphi_{2,1} - b_2}{r + \lambda_1} - \frac{\varphi_{1,1} - b_1}{r + \lambda_1}\right)$$

Now, conditional on receiving an offer, a type 2 in island 2:

- is the only applicant with probability $e^{-\nu_2 q_2 (1-x_2)}$, and receives $\beta \frac{\varphi_{2,2} - b_2}{r+\lambda_2}$.
- is in competition with other type 2 workers with probability $1 - e^{-\nu_2 q_2 (1-x_2)}$, and receives 0 surplus.

$$b_2 = rU_2 = b + \nu_2 e^{-\nu_2 q_2 (1-x_2)} \beta \frac{\varphi_{2,2} - b_2}{r + \lambda_2}.$$

Combining the two value equations, the no-arbitrage condition $rU_1 = rU_2$ implies

$$\nu_1 e^{-\nu_1 q_1 x_2 h_2} e^{-\nu_1 q_1} \beta \frac{\varphi_{2,1} - b_2}{r + \lambda_1} + \nu_1 e^{-\nu_1 q_1 x_2 h_2} \left(1 - e^{-\nu_1 q_1}\right) \beta \left(\frac{\varphi_{2,1} - b_2}{r + \lambda_1} - \frac{\varphi_{1,1} - b_1}{r + \lambda_1}\right) = \nu_2 e^{-\nu_2 q_2 (1 - x_2)} \beta \frac{\varphi_{2,2} - b_2}{r + \lambda_2}$$

B.2 Job creation condition

Let V_j denote the value of having an open vacancy in island j. The no-profit condition imposes $V_j = 0$.

For island 2, $V_2 = 0$ implies:

$$\begin{aligned} 0 = rV_2 = & \left(1 - e^{-\nu_2 (1 - x_2) q_2} - (1 - x_2) \nu_2 q_2 e^{-(1 - x_2) \nu_2 q_2}\right) \frac{\varphi_{2,2} - b_2}{r + \lambda_2} \\ & + (1 - x_2) \nu_2 q_2 e^{-(1 - x_2) \nu_2 q_2} (1 - \beta) \frac{\varphi_{2,2} - b_2}{r + \lambda_2} \\ & - c_2 \end{aligned}$$

where the first line represents meetings with more than 1 applicant and the second line meetings with 1 applicant.

For island 1, $V_1 = 0$ implies:

$$\begin{aligned} 0 = rV_1 = & e^{-\nu_1 q_1 x_2 h_2} e^{-\nu_1 q_1} \nu_1 q_1 (1 - \beta) \frac{\varphi_{1,1} - b_1}{r + \lambda_1} \\ & + e^{-\nu_1 q_1 x_2 h_2} (1 - e^{-\nu_1 q_1} - e^{-\nu_1 q_1} \nu_1 q_1) \frac{\varphi_{1,1} - b_1}{r + \lambda_1} \\ & + e^{-\nu_1 q_1 x_2 h_2} e^{-\nu_1 q_1} \nu_1 q_1 x_2 h_2 (1 - \beta) \frac{\varphi_{2,1} - b_2}{r + \lambda_1} \\ & + (1 - e^{-\nu_1 q_1 x_2 h_2} - e^{-\nu_1 q_1 x_2 h_2} \nu_1 q_1 x_2 h_2) \frac{\varphi_{2,1} - b_2}{r + \lambda_1} \\ & \nu_1 q_1 x_2 h_2 e^{-\nu_1 q_1 x_2 h_2} (1 - e^{-\nu_1 q_1}) \left((1 - \beta) \frac{\varphi_{2,1} - b_2}{r + \lambda_1} + \beta \frac{\varphi_{1,1} - b_1}{r + \lambda_1}\right) \\ & - c_1 \end{aligned}$$

B.3 Calibration

In addition to the calibration of standard parameters described in the main text, calibrating the model also requires parameter values for the output of each type of match; $\varphi_{1,1}$, $\varphi_{2,1}$, $\varphi_{2,2}$. Normalizing the mean level of $\varphi_{2,2}$ to one, we pin down the mean levels of $\varphi_{1,1}$ and $\varphi_{2,1}$ from the average measured under-employment rate and from the average wage premium earned by over-educated workers ($\frac{\omega_{2,1}}{\omega_{1,1}}$).

The vacancy posting costs are set to match the average job openings to job seekers ratio in typical college-level occupations (professional, services) and in typical high-school level occupations (construction and sales) as measured from the Help Wanted OnLine (HWOL) dataset from the Conference Board. Indeed, while job openings by degree-requirements is not observed, we do observe job openings by occupation, and while occupation groups are different from degree requirements, for the four high-level occupation groups (professional, services, construction and sales), there is a relatively straightforward mapping between occupation and degree requirements: services and professional occupations require a bachelor's degree in most cases (it is true for more than 90 percent of the 3-digit occupations in those two categories), while construction and sales typically require less than a high school degree (it is true for more than 75 percent of the 3-digit occupations in those two categories). We thus proxy job openings with high degree-requirements with job openings in services and professional occupations, and we proxy job openings with low degree-requirements with job openings in construction and sales.

To use the urn-ball matching function in a quantitative context, we proceed as Blanchard and Diamond (1994) and introduce a matching efficiency term. We set matching efficiency in each island to match the average job finding rates of respectively high-educated and low-educated workers. Specifically, we assume that workers send out an application with probability ν. In that case, it is easy to show that a worker's job finding rate searching in some homogenous island with queue length q is given by $\nu f(\nu q)$ with

$$f(q) = \frac{1 - e^{-q}}{q}$$

where ν denotes the "matching efficiency" of the labor market.[37] We allow ν to vary across islands with ν_j the matching efficiency of island j.

The observable job finding rates of type-2 and type-1 workers, $f_{2,t}$ and $f_{1,t}$, satisfy

$$\begin{cases} f_{2,t} = (1 - x_{2,t})\, f_{22,t} + x_{2,t} f_{21,t} \\ f_{1,t} = f_{11,t} \end{cases} \tag{7}$$

with $f_{22,t}$, $f_{21,t}$, $f_{11,t}$ denoting respectively the job finding rate of type 2 workers searching in island 2, the job finding rate of type 2 workers searching in island 1,[38]

[37]With q the ratio of job seekers to vacancies, the number of hires is given by $V(1 - e^{-q})$, so that the probability that a worker who sends one application gets a job is $f(q) = \frac{1-e^{-q}}{q}$.

[38]Note that, while $f_{2,t}$ and $f_{1,t}$ are observed, we do not observe $f_{22,t}$ and $f_{21,t}$, since we do not know where high-skill workers are searching.

and the job finding rate of type 1 workers searching in island 1. It is easy to show that $f_{22,t}$, $f_{21,t}$ and $f_{11,t}$ are given by:[39]

$$\begin{cases} f_{2,2} = \nu_2 f(\nu_2 q_2(1 - x_2)) \\ f_{2,1} = \nu_1 f(\nu_1 q_1 h_2 x_2) \\ f_{1,1} = \nu_1 e^{-\nu_1 q_1 h_2 x_2} f(\nu_1 q_1) \end{cases} \tag{8}$$

B.4 Underemployment rate: flow versus stock

In a dynamic setting, there is a difference between $x_{2,t}$, the (unobserved) fraction of high-skill workers *searching* in the low-tech island, and the under-employment rate, UE_t, which is the (observed) fraction of high-skill workers *employed* in the low-tech island and is given by

$$UE_t = \frac{N_{21}(t)}{N_{21}(t) + N_{22}(t)} \tag{9}$$

where $N_{21}(t)$ is the number of high-skill workers (type-2) employed in island 1 at time t, and $N_{22}(t)$ is the number of high-skill workers employed in island 2 at time t.

To link $x_{2,t}$ and UE_t, we can use a stock-flow accounting model of under-employment. Specifically, we consider a continuous time environment in which data are available only at discrete dates. For $t \in \{0, 1, 2...\}$, we refer to the interval $[t, t + 1[$ as 'period t'. We assume that during period t, the fraction of high-skill workers searching in island 1 is constant at $x_{2,t}$, the ratio of high-skill to low-skill job seekers is constant at $h_{2,t}$ and that the job finding rates are constant at $f_{22,t}$, $f_{21,t}$ and $f_{11,t}$.

Over $\tau \in [0, 1]$, the number of high-skill workers employed in island 2, $N_{22}(t+\tau)$, the number of high-skill workers employed in island 1, and $N_{21}(t + \tau)$, and the number of low-skill workers employed in island 1, $N_{11}(t + \tau)$, evolve according to

$$\begin{cases} \frac{dN_{22}(t+\tau)}{d\tau} = (L_2 - N_{22}(t + \tau) - N_{21}(t + \tau))\,(1 - x_{2,t})f_{22,t} - \lambda_{22} N_{22}(t + \tau) \\ \frac{dN_{21}(t+\tau)}{d\tau} = (L_2 - N_{22}(t + \tau) - N_{21}(t + \tau))\, x_{2,t} f_{21,t} - \lambda_{21} N_{21}(t + \tau) \\ \frac{dN_{11}(t+\tau)}{d\tau} = (L_1 - N_{11}(t + \tau))\, f_{11,t} - \lambda_{11} N_{11}(t + \tau) \end{cases} \tag{10}$$

with the job finding rates in each island given by (8).

With the observed under-employment rate UE_t given by (9), we can combine (10) with expressions for the job finding rates (8), and solve the system of differential

[39] To obtain (8) note that the number of high-type job seekers in island 2 is $(1 - x_{2,t})L_{2,t}$ and that the number of high-type job seekers in island 1 is $x_{2,t} h_{2,t} L_{1,t}$. For low-skill workers, we get a similar expression, except that they only get a job in they face no high-skill applicants, which happens with probability $e^{-\nu_1 x_{2,t} h_{2,t} q_{1,t}}$.

equation to express $\hat{x}_{2,t}$ as a function of $x_{2,t}$, $h_{2,t}$, $q_{2,t}$, $q_{1,t}$, $\hat{x}_{2,t-1}$ and the other constant parameters of the model.