Mendolicchio, Concetta; Paolini, Dimitri; Pietra, Tito

Working Paper

Investments in education and welfare in a two-sector, random matching economy

IAB-Discussion Paper, No. 8/2011

Provided in Cooperation with:
Institute for Employment Research (IAB)

Suggested Citation: Mendolicchio, Concetta; Paolini, Dimitri; Pietra, Tito (2011) : Investments in education and welfare in a two-sector, random matching economy, IAB-Discussion Paper, No. 8/2011, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg

This Version is available at:
https://hdl.handle.net/10419/57460

Institute for Employment Research
The Research Institute of the Federal Employment Agency

IAB-Discussion Paper 8/2011

Articles on labour market issues

Investments in education and welfare in a two-sector, random matching economy

Concetta Mendolicchio
Dimitri Paolini
Tito Pietra

Investments in education and welfare in a two-sector, random matching economy

Concetta Mendolicchio (IAB)
Dimitri Paolini (DEIR and CRENoS, Università di Sassari)
Tito Pietra (DSE, Università di Bologna)

Contents

Abstract

We consider a random matching model where heterogeneous agents choose optimally to invest time and real resources in education. Generically, there is a steady state equilibrium, where some agents, but not all of them, invest. Regular steady state equilibria are constrained inefficient in a strong sense. The Hosios (1990) condition is neither necessary, nor sufficient, for constrained efficiency. We also provide restrictions on the fundamentals sufficient to guarantee that equilibria are characterized by overeducation (or undereducation), present some results on their comparative statics properties, and discuss the nature of welfare improving policies.

Zusammenfassung

Im Rahmen eines Matching-Modells wird die optimale Investitionsentscheidung heterogener Agenten in Humankapital untersucht. Generisch existiert ein Gleichgewicht, in dem einige, aber nicht alle Agenten investieren. Reguläre Steady-State-Gleichgewichte sind im starken Sinne beschränkt ineffizient. Die Hosios-Bedingung ist für beschränkte Effizienz weder notwendig noch hinreichend. Weiter zeigen wir Restriktionen auf, die hinreichend sind für die Existenz eines Gleichgewichts, das durch Über- oder Unterinvestition in Bildung gekennzeichnet ist. Schließlich diskutieren wir komparativ-statische Eigenschaften des Modells sowie wohlfahrtssteigernde Politikmaßnahmen.

JEL classification: J24; H2

Keywords: Human capital; Efficiency; Labour Supply

Acknowledgements: The paper is part of the Ph.D. dissertation at IRES, Universitè Catholique de Louvain, of the first author. We thank for helpful comments the members of her dissertations committee (R. Boucchekine, V. Vandenberghe, M. Belot, B. Decreuse, B. Van der Linden), and participants at the 16^{th} Ecole de Printemps, Aix-En-Provence 2007, E.S.P.E. 2008, and at seminars at the Universities of Bologna, Essex, and Statale of Milan. Comments by A. Giovagnoli and P. Siconolfi have been helpful. The usual disclaimers apply. The first author acknowledges the support of the "Programma visiting professor" of the Università di Sassari (Italy). We acknowledge the financial support of MIUR, PRIN 2006, and of the Fondazione Banco di Sardegna.

1. Introduction

Investments in human capital and their welfare implications have been a central issue in economics for quite a few years. Still, the analysis is far from settled both empirically and theoretically. We study the welfare properties of equilibrium investments in human capital in a frictional economy. Specifically, we consider a two-sector version of the canonical Pissarides-Mortensen model (see, for instance, Pissarides (2000)), with random matching and two different markets, for qualified and non-qualified labor. The basic structure of the economy is simple. At birth, agents optimally choose wether or not to invest a fixed amount of time and real resources in education to get the opportunity to enter, after graduation, the qualified labor market. The two types of labor are hired by two separated group of firms. Wages are determined by Nash bargaining, after a random match occurs. Therefore, we basically embed a Roy (1951) model of investments in human capital into a two-sector random matching model.

Several previous contributions have studied investments in human capital in frictional labor markets, and, in particular, in random matching models. Among many others, Laing, Palivos and Wang (1995), Acemoglu (1996), Burdett and Smith (1996, 2002), Booth et alii (2005, 2006 and 2007). All these papers consider economies with a unique labor market and where investments in human capital increase the number of efficiency units of labor associated with a (physical) unit of time. Closer to our set-up are Sattinger (2003), Charlot and Decreuse (2005, 2006), Charlot, Decreuse and Granier (2005), and Tawara (2007), which consider economies a la Roy. Bear in mind that the two approaches to the analysis of investments in human capital (efficiency units versus heterogeneity of jobs and binary choice) emphasize different phenomena, and may have quite different welfare implications, because the efficiency units approach rules out, by assumption, the potentially important effects of the self-selection of workers into different labor markets. They also have significantly different empirical implications, see, for instance, Carneiro, Heckman, and Vytlacil (2001), and Cunha and Heckman (2006).

Our model is related to the contributions mentioned above, and, mostly, to Charlot and Decreuse (2005). In our, as in their, economy, the choice to invest in education is binary, and human capital acquired through schooling is not fungible.[1] However, we consider a class of economies with several new features. First, workers are heterogeneous along several different dimensions: productivity on the job (and unemployment benefits, or home production, if out of work) as qualified and unqualified, and probability of graduation. We do not to impose any restriction on the correlations across these variables, allowing different agents to have comparative advantages in different jobs, which is consistent, for instance, with the results in Cunha, Heckman, and Navarro (2005). Secondly, schooling has direct costs and takes time. Time to educate is an important phenomenon, because, empirically, the related costs are, by large, the most important component of the total costs of education.[2] Therefore, it is worthwhile to consider them explicitly, as standard in the literature (see, Becker (1964), Ben-Porath (1967), and most of the subsequent studies). In our model, total costs are endogenous, determined by unemployment rates and wages (and by the direct costs). Third, we consider two separate labor markets with different productivities and (potentially) different vacancy costs and matching functions, so that unemployment rates may vary across levels of education, which is consistent with a large empirical evidence. Variations in the unemployment rates are determined by differences in the

[1] At least in developed countries, this is best interpreted as a choice about attending college or a professional school. For the US, it is maybe better to think about graduate degrees. This choice still involves a significant share of the population. According to Snyder, Dillow, and Hoffman (2009, Table 268), in 2005-06, were awarded around 1.49 milion B.A. degrees (4 years college only) and 0.68 milion Master and First professional Degrees.

[2] For instance, in Western Europe, they are usually over 90% of the total costs and, in several countries, direct costs are actually negative, according to estimates reported in de la Fuente (2004).

"labor market institutions" variables and, loosely speaking, in the conditional expectation of the productivities in the two markets. Finally, we assume that, when agents invest in education, they fail to graduate with some positive probability.[3]

Throughout the paper, we just consider steady states, and exploit the steady state condition in the definition of several features of the economy, and of the behavioral functions. Section 2 describes the structure of the model. In Section 3, we show that, under appropriate assumptions, and generically in the space of the economies[4], there are interior steady state equilibria (SSE), i.e., SSE where some, but not all, the agents invest in education, provided that the length of schooling, denoted T, is not "too large". Moreover, we show that interior SSE are generically regular. All the results in this section are mainly of a technical nature, and preliminary to our analysis of efficiency. Therefore, most of their cumbersome details are in Appendix 7.3.

Section 4 is the core of the paper. We propose a notion of constrained efficiency, WCO, based on the idea that it should not be possible to improve upon the market allocation, in terms of expected total surplus, by simply modifying the set of people getting educated and letting the endogenous variables adjust to their new equilibrium values. In our class of economies, there are three possible sources of inefficiency: congestion in the labour markets (as in the usual random-matching models), irreversibility of investments leading to a possible hold-up problem (as in Acemoglu (1996)), self-selection into separate labor markets (as in Charlot and Decreuse (2005)). Our general set-up allows us to study the interaction of these three phenomena. The main result is that, typically, interior SSE are not weakly constrained efficient. At a SSE, both overinvestment in education, as in Charlot and Decreuse (2005), or underinvestment, as in Acemoglu (1996) are possible. Moreover, contrary to the standard results in one-sector random matching models, the Hosios (1990) condition[5] is neither necessary, nor sufficient, for WCO.

The aim of the paper is purely theoretical. Random matching models and Roy models are widely used in the literature. Therefore, it is interesting to study, in a general setting, the equilibrium outcomes of an economy presenting both features. We adopt a set of restrictions on the parameters of the economy as weak as possible. Evidently, there is a trade off here. Some of our results are less sharp than the ones obtainable in more parametric exercises, as the one carried out in Charlot and Decreuse (2005). On the other hand, drastically simplifying assumptions tend to generate models with clearly counterfactual implications.[6] In Section 5, we move on to a simplified version of our economy, and focus on the two cases of complementarity and substitutability between ability and education. We provide sufficient conditions for overeducation (in the case of complementarity) and undereducation (with substitutability). For the same two cases, we also establish the comparative statics properties of interior SSE. These properties are different in the two cases, so that, in principle, one could use them to test, empirically, for overeducation vs. undereducation. Moreover, under the restrictions which allow us to identify the comparative statics properties of SSE and the nature of inefficiency, we establish that, to improve welfare, it suffices to introduce small taxes (or subsidies) on the direct costs of education.

While our assumptions on the parameters describing the economy are weak, the structure of

[3] This is consistent with the data: for instance, in Western Europe, the college drop-out rate varies between 15% in Ireland and 58% in Italy (See OECD (2004)). However, one may wonder why to introduce individual risk in a model with risk-neutral agents. Apart from descriptive realism, this assumption simplifies the proof of the existence of steady state equilibria, and has no substantive effect on our other results (in the general setting). We conjecture that they could all be established without it.

[4] Our space of the economies is infinite dimensional. Therefore, as usual in the literature, a property is generic if it holds for an open, dense set of economies.

[5] The condition is that, at the SSE, the absolute value of the elasticity of the matching function is equal to the workers' weight in the bargaining process determining their wage. In the basic, one-sector, random matching model, this is a necessary and sufficient condition for constrained efficiency of SSE.

[6] For instance, in their model, given that the production functions are linear in innate ability, the wage premium is constant across workers.

the model incorporates several very strong assumptions. For instance, there is no on-the-job, or while-in-school, search; even if unemployed, workers cannot simultaneously search in the two different markets; heterogeneity is just one–sided, because firms are identical; labor supply and investments in physical capital are perfectly inelastic;[7] after schooling, if there is any, all the agents are in the labor force, either employed or searching for a job. Evidently, relaxing some of these assumptions could have relevant implications (see, for instance, Shimer (2006), Burdett, Carrillo-Tudela, and Coles (2009), and Booth et alii (2005, 2006, 2007)).

Finally, and maybe most important, we consider a random matching model. In the last decade, the directed search approach has become increasingly common in the literature. Its key feature is that the wage setting mechanism induces agents to internalize the congestion externalities, so that the equilibrium outcome may be constrained efficient. It is quite reasonable to think that an appropriately designed wage profile could take care simultaneously of all the different sources of constrained inefficiency at work in our class of economies. However, in one-sector models, the source of inefficiency is just related to a congestion externality which can be characterized using the Hosios's condition. This is not true here. Therefore, the results obtained in one-sector models do not immediately generalize to a two-sectors model with self-selection. The extension of the analysis to take into account these important, additional, phenomena is an open issue, left for future research.[8]

2. The Model

We start discussing the demographic structure of the economy. Time is continuous. There is a continuum of agents, denoted by θ_i, where $\theta \in \Theta^0 = [\theta_\ell, \theta_h]$ describes the agent's innate characteristics, while $i \in [0,1]$ is just used to index otherwise identical agents. The Lebesgue measure of any set $A \subset \mathbb{R}^M$, for some M, is denoted $\mu(A)$. We endow Θ^0 and $[0,1]$ with the Lebesgue measure, normalized so that $\mu(\Theta^0) = \mu([0,1]) = 1$, and $\Theta = \Theta^0 \times [0,1]$ with the product measure.

As usual, we assume that workers drop out of the economy ("ie"). If agent θ_i dies, he is replaced by his own clone, so that the distribution of the agents across θ and the measure of the agents are stationary. The dynamics of the population is deterministic at the aggregate level, but follows a stochastic process at the individual level. Let $\widehat{\theta}(t)$ be the set of type θ agents alive at t and born at $\tau < t$, and let $\mu(\widetilde{\theta}(t))$ be its measure. We start imposing

Assumption 1 (A1)

1. For each agent θ_i, death occurs according to a Poisson process with arrival rate γ;
2. For each t, and each θ, the rate of change of the measure of the set of agents of type θ due to death is $-\gamma$.

These two properties are usually jointly justified assuming (1), and that the random variables are i.i.d.. Then, one derives (2) by what Shimer (2005, p. 1001) refers to as "a standard abuse of the law of large numbers" (LLN in the sequel). In Appendix 7.1, we adapt to our set-up a well-known construction which properly justifies A1, and the analogous assumptions (A2 and A4) introduced later on. The construction is based on the assumption that the realizations of

[7] An analysis of a (static) Roy model with labor market imperfections, and elastic investments in human and physical capital is in a companion paper (Mendolicchio, Paolini, and Pietra (2008)).

[8] We thank, in particular, B. Decreuse for stimulating comments and suggestions on these issues.

the individual random variables are correlated across i, for each θ, i.e., on the violation of the i.i.d. assumption.

At birth, each individual is uneducated, or unskilled, denoted by a superscript $k = ne$. By spending a period of fixed length T in education, and paying direct flow costs described by a function $\widehat{c}(\theta)$, he becomes educated (or "qualified", or "skilled"), denoted by a superscript $k = e$, with probability $\alpha(\theta)$. For instance, agent θ_i born at t may choose to enter immediately the unskilled labor force, or to postpone his participation to the labor market at time $(t+T)$. If so, he goes to school for a period of length T. At time $(t+T)$, the random variable "success/failure in education" realizes. If θ_i graduates (this realizes with probability $\alpha(\theta)$), he may enter the skilled labor force. Otherwise, he enters the unskilled one. Therefore, at each t, there are three sets of individuals: qualified workers, unqualified workers, and students.

For each θ, the individual random variable success/failure in education realizes at the end of schooling. This is a strong assumption (and generally inaccurate from the descriptive viewpoint). However, as long as the probability of failure is exogenous, our results are robust to more realistic descriptions of the phenomenon.

Assumption 2 (A2):

1. For each agent θ_i born at τ, the probability of graduation (at $(\tau+T)$, conditional of having invested in education and being alive at $(\tau+T)$) is $\alpha(\theta)$;
2. For each θ, $\alpha(\theta)$ is the fraction of the cohort born in the time interval $(\tau, \tau+\Delta)$ who actually graduates (conditional on their having invested in education and being alive in the interval $(\tau+T, t+T+\Delta)$), for each Δ.

For notational convenience, we rename agents so that, at the steady state, for each θ, agents who actually graduate have $i \in \left[0, e^{-\gamma T}\alpha(\theta)\right]$.[9] Given that type θ individual are identical, there is no loss of generality in this notational convention.

Productivities on the job and in home production (and/or unemployment benefits) depend upon innate characteristics and the level of education. If educated and working as such, a worker has output $f^e(\theta)$. If unemployed, he has home production $b^e(\theta)$. Otherwise, he produces $f^{ne}(\theta)$ (or $b^{ne}(\theta)$). We assume that, after graduation, workers cannot simultaneously search for a job in both markets. Hence, to simplify notation (and, at a SSE, without any loss of generality, because education is costly), educated individuals only look for a job on market e. The functions (f, b) are time-invariant. Hence, human capital does not depreciate, and there is no learning-by-doing. More realistic assumptions would not affect the results, as long as these additional phenomena are described by exogenous processes.

More formally, instantaneous output is given by a function $f^k : [\theta_\ell, \theta_h] \times \{e, ne\} \to \mathbb{R}_+$. Home production by a function $b^k : [\theta_\ell, \theta_h] \times \{e, ne\} \to \mathbb{R}_+$. We assume that all the relevant functions are at least C^3 on some open neighborhood $(\theta_\ell - \varepsilon, \theta_h + \varepsilon)$, that individuals are more productive when educated (i.e., that $f^e(\theta) > f^{ne}(\theta)$, and $b^e(\theta) \geq b^{ne}(\theta)$, for each θ), and that productivity on-the-job is larger than home production. These are fairly natural assumptions. Moreover, we label individuals so that $f^e(\theta)$ is strictly increasing. Bear in mind that θ is merely an ordering parameter, because no restriction (but continuous differentiability) is imposed on the other relevant functions, and consequently, for instance, on the expected gains in productivity

[9] If a set of agents indexed by $[0, 1]$ is born at time t and invest in education, a measure $e^{-\gamma T}$ of agents is still alive at time $(t+T)$ and a fraction $\alpha(\theta)$ of them graduates.

due to education. Hence, we are completely agnostic regarding the existence of some intrinsic characteristic (say, "innate ability") of the individuals inducing systematic correlations between their performances in different activities.[10] In particular, we allow different agents to have comparative advantages in different jobs, which is consistent, for instance, with the results in Cunha, Heckman, and Navarro (2005). We just introduce the (very mild) assumption that, for every agent, the productivity on the job is larger when educated.

To summarize,

Assumption 3 (A3):

1. For each k, $f^k(\theta), b^k(\theta), \alpha(\theta), \widehat{c}(\theta) \in C^3$ on $(\theta_\ell - \varepsilon, \theta_h + \varepsilon)$, for some $\varepsilon > 0$;
2. $f^e(\theta)$ is strictly monotonically increasing in θ on the set $(\theta_\ell - \varepsilon, \theta_h + \varepsilon)$;
3. For each k, $(f^k(\theta) - b^k(\theta))$ is bounded above and bounded away from 0, $1 > \alpha(\theta) > 0$, and $\widehat{c}(\theta) > 0$, for each θ;
4. For each θ, $f^e(\theta) > f^{ne}(\theta)$ and $b^e(\theta) \geq b^{ne}(\theta)$.

Agents are endowed with a Von Neumann - Morgenstern utility function and are risk-neutral.[11] Hence, there is no essential loss of generality in assuming that each agent knows his own type θ. A firm, after the match, observes the value θ of the agent it is matched with (i.e., it observes $f^k(\theta)$ and $b^k(\theta)$, the only relevant variables from its viewpoint).

A final remark on notation: We will often take integrals of some function of θ, say, for instance, $f^e(\theta)$, over some (for now, arbitrary) set $\Omega \subset \Theta$. To avoid confusion, we will use the notation $\int_\Omega f^e(\vartheta)\, d\vartheta$ and use, for instance, $\frac{\partial \int_\Omega f^e(\vartheta) d\vartheta}{\partial \theta}|_{\theta=\theta_m}$ to denote the derivative with respect to the bound θ_m of (one of the) intervals of integration (assuming that this is meaningful). Also to simplify the notation, the same (arbitrary) function $V(.)$ will be sometime written $V(x, y)$, for some pair of vectors (x, y), sometimes $V(x)$ or $V_y(x)$. This simply means that the "missing" vector y is taken as fixed. Finally, we are only interested in steady states, and we will omit the index t, whenever possible.

2.1. Workers

Existence of a continuum of agents with identical θ is introduced just for technical reasons, i.e., to guarantee that there are always workers active in sector ne. In the sequel, whenever possible, we will omit the subscript "i", and leave implicit the "almost surely" qualification of many of our statements.

After birth (or after attending school) agents enter the labor market, searching for a job. An agent, if active on labor market k, receives job offers according to a Poisson process with arrival rate π^k, endogenously given by a (possibly, $k-$dependent) matching function. Let $\pi \equiv (\pi^{ne}, \pi^e)$.

[10] This is also because, to obtain significantly stronger conclusions, or a substantive simplification of the argument, we would need much more than a simple restriction on the signs of the derivatives, see Section 5.

[11] In the one-sector model, too, agents maximize their discounted expected income (instead of expected utility of income). This can be justified making appeal to market completeness. Here, we would need market completeness with respect to the risk of failure in education, a much less compelling assumption, because of possible moral hazard problems. On the other hand, abstracting from them, the main results of the paper could be established for risk-averse individuals, provided that the degree of risk aversion is sufficiently small.

Let r' be the (type- and education-invariant) interest rate, and $w^k(\theta,\pi)$ be the wage rate of a type θ agent, if k. Let $(\gamma + r') \equiv r$.

Define as $U^k(\theta,\pi)$ the expected life-time utility of search of an agent of type θ and education k, as $V^k(\theta,\pi) = \frac{w^k(\theta,\pi)}{r}$ his expected life-time utility of a match. As usual, $U^k(\theta,\pi) = \frac{\pi^k w^k(\theta) + r b^k(\theta)}{r(r+\pi^k)}$.

Assume that capital markets are perfect and, without any essential loss of generality, let $\widehat{c}(\theta)$ be time-invariant. At birth, the discounted, expected utility of an agent of type θ investing in education is

$$H(\theta,\pi;T) = e^{-rT}\left[\alpha(\theta)U^e(\theta,\pi) + (1-\alpha(\theta))U^{ne}(\theta,\pi)\right] - \frac{\left(1-e^{-rT}\right)}{r}\widehat{c}(\theta)\,,$$

i.e., if an agent chooses to invest in education, he bears the direct costs up to period T. Then, if he graduates (which happens with probability $\alpha(\theta)$), he enters the labor market for educated workers. Otherwise (with probability $(1-\alpha(\theta))$, he enters the other market.

Evidently, an agent invests in education only if $H(\theta,\pi;T) \geq U^{ne}(\theta,\pi)$. Solving $H(\theta,\pi;T) - U^{ne}(\theta,\pi) = 0$, and using continuity of the maps, we can partition Θ^0 into two (measurable) subsets of types of agents: the ones investing in education and the others. For the sake of concreteness, let's assume that all the indifferent agents choose to invest. Hence, by our tie-breaking rule, rearranging variables, and multiplying by e^{rT}, an agent of type θ chooses to get educated if and only if

$$G(\theta,\pi;T) \equiv \alpha(\theta)\left(U^e(\theta,\pi) - U^{ne}(\theta,\pi)\right) + \left(1-e^{rT}\right)\left(c(\theta) + U^{ne}(\theta,\pi)\right) \geq 0, \tag{1}$$

where $c(\theta) \equiv \frac{\widehat{c}(\theta)}{r}$. The function $e^{-rT}G(\theta,\pi;T)$ is the discounted, expected value of the investment in education.

Assuming stationarity, define $\Theta^{0k}(\theta,\pi;T)$, for $k = e, ne$, the set of types investing (if $k = e$) or not investing (if $k = ne$) in education. The two sets are implicitly defined by the conditions $G(\theta,\pi;T) \geq 0$ and $G(\theta,\pi;T) < 0$, respectively. Using A1-A2, define the correspondences describing (modulo a relabelling of the agents) the stationary labor force:

a. $\Theta^{e\alpha}(\pi;T) = \left\{\theta_i \in \Theta | G(.) \geq 0 \text{ and } i \leq e^{-\gamma T}\alpha(\theta)\right\}$, the set of educated workers,

b. $\Theta^{ne\alpha}(\pi;T) = \left\{\theta_i \in \Theta | G(.) < 0 \text{ or } G(.) \geq 0 \text{ and } e^{-\gamma T} \geq i > e^{-\gamma T}\alpha(\theta)\right\}$, the set of uneducated workers.

Evidently,

$$\begin{aligned}
\mu\left(\Theta^{e\alpha}(\pi;T)\right) &= e^{-\gamma T}\int_{\Theta^{0e}(\pi;T)} \alpha(\vartheta)\,d\vartheta \\
&\quad \text{and} \\
\mu\left(\Theta^{ne\alpha}(\pi;T)\right) &= \int_{\Theta^{0ne}(\pi;T)} d\vartheta + e^{-\gamma T}\int_{\Theta^{0e}(\pi;T)} (1-\alpha(\vartheta))\,d\vartheta \\
&= \mu\left(\Theta^{0ne}(.)\right) + e^{-\gamma T}\left(\mu\left(\Theta^{0e}(.)\right) - \mu\left(\Theta^{e\alpha}(.)\right)\right).
\end{aligned} \tag{2}$$

By continuity of $G(\theta,\pi;T)$, $\Theta^{0e}(\pi;T)$ is the union of a finite collection of closed intervals and (possibly) of isolated points, at any (π,T).

2.2. Firms

Each firm is endowed with a technology allowing it to use one unit of labor to produce a quantity of output. Given that, at any SSE, expected profits are nil, to restrict firms to be active only in one market does not entail any loss of generality. Firms (potentially) active in each sector are identical, and there is an unlimited number of potential entrants in each industry. We can think of the two sectors as defined by different technologies used to produce the same good, or (more plausibly) as sectors producing different commodities. Without any loss of generality, both output prices are equal to 1, so that we can ignore them.

The set-up of the demand side of each one of the two labor markets is standard. Under perfect capital markets, firms have the same discount factor r'. To open a vacancy in labor market k entails a fixed, flow cost, v^k, for $k = ne, e$, satisfying $v \equiv (v^{ne}, v^e) >> 0$. Usually, they are interpreted as advertising and recruitment costs. Also, remember that γ is the arrival rate of the Poisson process describing the dissolution of a match and that $(\gamma + r') \equiv r$. Finally, for each firm active in market k, matches are governed by a Poisson process with arrival rate q^k.

The flow of profits induced by a vacancy filled by a θ worker is $\left(f^k(\theta) - w^k(\theta)\right)$, until he drops out of the match. Hence, its expected, discounted value (conditional on $\Theta^{k\alpha}(\pi; T)$) is

$$J^k = \frac{\int_{\Theta^{k\alpha}(\pi;T)} \left(f^k(\vartheta) - w^k(\vartheta, \pi)\right) d\vartheta}{r\mu\left(\Theta^{k\alpha}(\pi; T)\right)}. \tag{3}$$

Vacancies in a market are created as long as their expected value is positive. Hence, assuming perfect competition, the zero expected profit condition requires

$$\frac{\int_{\Theta^{k\alpha}(\pi;T)} \left(f^k(\vartheta) - w^k(\vartheta, \pi)\right) d\vartheta}{r\mu\left(\Theta^{k\alpha}(\pi; T)\right)} = \frac{v^k}{q^k}. \tag{4}$$

2.3. Bargaining

After each match, the shares of output of worker and firm are determined by a bargaining process, taking place after the type of the worker is revealed (equivalently, the wage is output - i.e., worker's type - contingent). The firm and the worker bargain over their shares of total output adopting the Nash bargaining solution criterion, with exogenous weights respectively $(1 - \beta)$ and β. The outside options are, respectively, $U^k(\theta, \pi)$, for a worker of type θ and qualification k, and 0 for each firm, by the free-entry assumption. Hence, each firm always hires the first worker it meets. The output shares are obtained solving

$$\max \left(\frac{w^k(\theta, \pi) - rU^k(\theta, \pi)}{r}\right)^{\beta} \left(\frac{f^k(\theta) - w^k(\theta, \pi)}{r}\right)^{1-\beta},$$

and, as usual, the wage of a $\theta-$worker active in sector k is

$$w^k(\theta, \pi) = \beta \frac{\left(r + \pi^k\right) f^k(\theta)}{r + \beta\pi^k} + (1 - \beta) \frac{rb^k(\theta)}{r + \beta\pi^k}. \tag{5}$$

2.4. Matching and unemployment

At a steady state, on market k, the measure of unemployed agents is $u^k\mu(\Theta^{k\alpha}(\pi; T))$, while $o^k\mu(\Theta^{k\alpha}(\pi; T))$ is the measure of vacant jobs ("openings"), expressed in terms of the measure

of the labor force of type k. Define the "market tightness" variable $\phi^k \equiv \frac{o^k}{u^k}$, for each k. Let

$$q^k = q^k(\phi) \equiv m^k(\frac{1}{\phi^k}, 1) \text{ and } \pi^k = \pi^k(\phi) \equiv \phi^k q^k(\phi) \equiv m(1, \phi^k),$$

be, for each k, the functions associating with each ϕ^k the arrival rate of the Poisson process governing the flows of job applications and job offers in sector k. As usual, we adopt

Assumption 4 (A4):

1. $m^k(u^k, o^k) \in C^3$ satisfies $\nabla m^k >> 0$, is concave, homogeneous of degree 1 in (u^k, o^k) and satisfies the Inada's condition;
2. $q^k(\phi) \in C^3$ satisfies $\frac{\partial q^k}{\partial \phi^k} < 0$, and $\frac{\partial q^k}{\partial \phi^{k'}} = 0$. Moreover, $\eta_{q^k}(\phi) \equiv \frac{\phi^k}{q^k(\phi)} \frac{\partial q^k}{\partial \phi^k} \in (-1, 0)$;
3. $\pi^k(\phi) \in C^3$ satisfies $\frac{\partial \pi^k}{\partial \phi^k} > 0$ and $\frac{\partial \pi^k}{\partial \phi^{k'}} = 0$. Moreover, $\eta_{\pi^k}(\phi) \equiv \frac{\phi^k}{\pi^k(\phi)} \frac{\partial \pi^k}{\partial \phi^k} = (1 + \eta_{q^k}(\phi)) \in (0, 1)$.[12]

Let $\widehat{\theta}^{k\alpha}(t)$ define the set of agents of type θ, alive at t, and active in labor market k. As in A1, we introduce the following:

Assumption 5 (A5):

1. For each θ and each $i \in \widehat{\theta}^{k\alpha}(t)$, if unemployed, job offers follow a Poisson process with arrival rate $\pi^k(\phi)$;
2. At each t, $\widehat{\theta}^{k\alpha}(t)$ and k, $\pi^k(\phi)$ is the rate of change, due to the transition into employment, of the measure of unemployed agents.

The size of the set of unemployed is affected by the flows of individuals dropping out of the labor force, at rate γ, or getting a job, at rate $\pi^k(\phi)$. The flow of individuals into unemployment is due to the new workers replacing the ones leaving the market. For the educated workers, at each t, it is given by the inflow of people who, at $(t - T)$, had chosen to get into education, and have both survived and succeeded, and has rate $\gamma\mu(\Theta^{e\alpha}(.))$. Therefore, at a stationary state,

$$\frac{\partial u^e \mu(\Theta^{e\alpha}(\pi; T))}{\partial t} = (-\gamma u^e - \pi^e(\phi) u^e + \gamma)\mu(\Theta^{e\alpha}(\pi; T)).$$

Using (2), the stationary rate of unemployment for the educated agents is

$$u^{e*} = \frac{\gamma}{\gamma + \pi^e(\phi)}. \tag{6}$$

Similarly, one can show that

$$u^{ne*} = \frac{\gamma}{\gamma + \pi^{ne}(\phi)}. \tag{7}$$

Evidently, these are the obvious translations in our setting of the usual stationary values of the rate of unemployment.

[12] Evidently, $\pi^k(\phi)$ (and $q^k(\phi)$) only depends upon ϕ^k, for each k. We write them this way to simplify notation.

2.5. The space of the economies

In the sequel, several results are established for a proper subset of economies. The parameters defining the economy are: vacancy costs, $v \in \mathbb{R}^2_{++}$, a pair (m^e, m^{ne}) satisfying A4, and a vector (f, b, c, α) satisfying A3. We always treat (γ, r'), hence r, as fixed. For convenience, we define the space of the economies directly in terms of $c(\theta) = r\widehat{c}(\theta)$. Thus, an economy is

$$\xi = (v, m, f, b, c, \alpha) \in \Xi.$$

We endow $\mathbb{R}^2_{++}$ with the Euclidean topology, all the functional spaces with the C^3 compact-open topology, and Ξ with the product topology. It is well known that Ξ is a metric space.

A property is generic if it holds for an open and dense subset of parameters. This is the standard notion of genericity for infinite dimensional spaces, given that there is no canonical notion of measure.

3. Equilibrium

Here, and in the sequel, we describe all the functions as depending upon ϕ, instead of π or q. Also, whenever required, we make explicit their dependence on (T, ξ).

At a steady state, replacing (5) into (1), we obtain

$$\begin{aligned} G(\theta, \phi; T, \xi) &= \alpha(\theta)\left[\frac{\pi^e(\phi)\beta f^e(\theta) + rb^e(\theta)}{r + \beta\pi^e(\phi)} - \frac{\pi^{ne}(\phi)\beta f^{ne}(\theta) + rb^{ne}(\theta)}{r + \beta\pi^{ne}(\phi)}\right] \\ &\quad + \left(1 - e^{rT}\right)\left[c(\theta) + \frac{\pi^{ne}(\phi)\beta f^{ne}(\theta) + rb^{ne}(\theta)}{r + \beta\pi^{ne}(\phi)}\right]. \end{aligned} \tag{8}$$

We define SSE in terms of the pair of "market tightness" variables $\phi = (\phi^e, \phi^{ne})$.

Definition 1. Given (T, ξ), an interior SSE is a pair $(\overline{\phi}^e, \overline{\phi}^{ne})$, and associated $(w^e(\theta, \overline{\phi}; T, \xi), w^{ne}(\theta, \overline{\phi}; T, \xi))$ and $(\Theta^{e\alpha}(\overline{\phi}; T, \xi), \Theta^{ne\alpha}(\overline{\phi}; T, \xi))$, such that:

i. $\frac{\int_{\Theta^{k\alpha}(\overline{\phi})}(f^k(\vartheta) - w^k(\vartheta, \overline{\phi}; T, \xi))d\vartheta}{r\mu(\Theta^{k\alpha}(\overline{\phi}))} = \frac{v^k}{q^k(\overline{\phi})}$, for each k,

ii. $w^k\left(\theta, \overline{\phi}; T, \xi\right) = \beta\frac{(r + \pi^k(\overline{\phi}))f^k(\theta)}{r + \beta\pi^k(\overline{\phi})} + (1 - \beta)\frac{rb^k(\theta)}{r + \beta\pi^k(\overline{\phi})}$, for each k,

iii. $\Theta^{0e}(\overline{\phi}; T, \xi) \neq \emptyset$ and $\Theta^{0e}(\overline{\phi}; T, \xi) \neq \Theta^0$.

Conditions $(i - ii)$ are standard, (iii) is the interiority condition.

Given that $\alpha(\theta) \in (0, 1)$, for each θ, the set $\Theta^{ne\alpha}(\phi; T, \xi)$ is always non-empty. On the contrary, it may very well be that $\Theta^{0e}(\phi; T, \xi) = \emptyset$ and, for T sufficiently large, the only SSE has this property. Anyhow, we will ignore these trivial equilibria.[13]

Replacing (ii) into (i), we can rewrite the zero-profit conditions as

$$\Phi^{k'}(\phi; T, \xi) \equiv \frac{1 - \beta}{r + \beta\pi^k(\phi^k)}\frac{\int_{\Theta^{k\alpha}(\phi; T, \xi)}(f^k(\vartheta) - b^k(\vartheta))d\vartheta}{\mu(\Theta^{k\alpha}(\phi; T, \xi))} - \frac{v^k}{q^k(\phi)} = 0, \text{ for each } k. \tag{9}$$

[13] Trivial "autarkic" equilibria with no vacancies and no labor force in one sector (or in both) also exist, as usual in random matching models. The difference is that, in our economy, for T sufficiently large, at each SSE, $\Theta^{0e}(\phi; T, \xi) = \emptyset$.

Let $\Phi'(\phi;T,\xi) \equiv (\Phi^{ne'}(\phi;T,\xi), \Phi^{e'}(\phi;T,\xi))$. Then, by definition, $\overline{\phi}$ is an interior SSE if and only if $\Phi'(\overline{\phi};T,\xi) = 0$, $\Theta^{0e}(\overline{\phi};T,\xi) \neq \Theta^0$, and $\Theta^{0ne}(\overline{\phi};T,\xi) \neq \emptyset$.

It is convenient to replace $\Phi'(\phi;T,\xi)$ with

$$\begin{aligned}\Phi(\phi;T,\xi) &\equiv \begin{bmatrix} \frac{r+\beta\pi^e(\phi)}{1-\beta}\Phi^{e'}(\phi;T,\xi) \\ \frac{r+\beta\pi^{ne}(\phi)}{1-\beta}\Phi^{ne'}(\phi;T,\xi) \end{bmatrix} \equiv \begin{bmatrix} F^e(\phi;T,\xi) - A^e(\phi^e;T,\xi) \\ F^{ne}(\phi;T,\xi) - A^{ne}(\phi^{ne};T,\xi) \end{bmatrix} \\ &\equiv \begin{bmatrix} \frac{\int_{\Theta^{e\alpha}(\phi;T,\xi)}(f^e(\vartheta)-b^e(\vartheta))d\vartheta}{\mu(\Theta^{e\alpha}(\phi;T,\xi))} - v^e\frac{r+\beta\pi^e(\phi^e)}{(1-\beta)q^e(\phi^e)} \\ \frac{\int_{\Theta^{ne\alpha}(\phi;T,\xi)}(f^{ne}(\vartheta)-b^{ne}(\vartheta))d\vartheta}{\mu(\Theta^{ne\alpha}(\phi;T,\xi))} - v^{ne}\frac{r+\beta\pi^{ne}(\phi^{ne})}{(1-\beta)q^{ne}(\phi^{ne})} \end{bmatrix}.\end{aligned}$$

Clearly, $\Phi'(\phi;T,\xi) = 0$ if and only if $\Phi(\phi;T,\xi) = 0$. Evidently, for each k, $F^k(\phi;T,\xi)$ depends upon ϕ only because of the effects of its changes on the set $\Theta^{k\alpha}(\phi;T,\xi)$, while, for each k, $A^k(\phi;T,\xi)$ does not depend upon $\Theta^{k\alpha}(\phi;T,\xi)$. This is very convenient in computations and therefore, in the sequel, we will mostly use this map.

There are two main difficulties in establishing the existence of interior SSE. First, for T large, there is always a trivial SSE where no one invests in education. Under very general (and natural) assumptions, for T sufficiently small, there is a SSE where all the agents invest in education. Therefore, the best we can look for is the existence of an interval (T_ξ, T^ξ) such that, for each $T \in (T_\xi, T^\xi)$, the economy ξ has an interior SSE. Secondly, we can use a fixed point argument to show the existence of SSE (once we impose some restrictions on the agents' expectations concerning the inactive sector, if any). However, fixed point arguments do not seem to suffice to establish the existence of interior SSE. Hence, we need a different approach. The proof is cumbersome and we postpone all the technicalities to Appendix 7.3. For it to work, we need an additional property, namely, that, at the unique SSE associated with the artificial economy where all the agents invest in education, if there are no costs of education, each agent would actually choose to invest. This is a very mild restriction, indeed. By continuity, we can formulate it as follows.

Assumption 6 (A6):

Let ϕ_ξ be the (unique) SSE of the artificial economy $\xi \in \Xi$ with $\Theta^{0e} = \Theta^0$. There is T_ξ such that $G(\theta, \phi_\xi; T, \xi) \geq 0$, for each $\theta \in \Theta^0$, if and only if $T \in (0, T_\xi]$.

Lemma A1 in Appendix 7.3.1 provides one set of restrictions on ξ such that A6 holds. However, A6 holds also under several other sets of assumptions and, most important, it is very natural in our set-up.

Theorem 1. Under A1-A6, for each $\xi \in \Xi^*$, an open, dense subset of Ξ, there is an interior SSE $\phi(T,\xi) \in C^1$ for each $T \in (T_\xi, T^\xi)$, an open subset of $(0,\infty)$.

Proof. See Appendix 7.3.1. ∎

The key point is that, modulo some arbitrarily small adjustment of production and education cost functions, and of vacancy costs, there is an interval (T_ξ, T^ξ) for which interior SSE exist. As already explained, this is the best result we can hope for, from the qualitative viewpoint.

Remark 1. The proof of Thm. 1 exploits appropriate perturbations of v, which are, possibly, $k-$dependent. They could be replaced by perturbations of $f^k(\theta) =$, for each k, and by local perturbations of the functions $q^k(\phi)$ (preserving, if so required, its invariance across markets).

Define the set

$$\Xi_T = \{\xi \in \Xi | \text{ given } T, \text{ there is an interior SSE, } \phi(T,\xi)\}.$$

Next, we study some properties of Ξ_T, for arbitrarily given $T > 0$. The results may be of some autonomous interest, but they are mainly motivated as a step to discuss (lack of) efficiency of SSE.

Theorem 2. For each $T > 0$, the set Ξ_T is non-empty and it contains an open set Ξ_T°. Moreover, Ξ_T° contains an open, dense subset Ξ_T^{reg} such that, for each $\overline{\xi} \in \Xi_T^{reg}$, at each interior SSE $\phi(T,\overline{\xi})$,
i. $G(\theta_\ell, \phi(\overline{\xi}); \overline{\xi}) \neq 0$ and $G(\theta_h, \phi(\overline{\xi}); \overline{\xi}) \neq 0$,
ii. $D_\phi \Phi(\phi; \overline{\xi})$ has full rank at $\phi = \phi(\overline{\xi})$,
iii. the number of interior SSE is finite, and there is an open neighborhood $V(\overline{\xi}) \subset \Xi_T^{reg}$ such that, at each $\xi \in V(\overline{\xi})$, interior SSE are locally described by a finite collection of C^1 maps, $(\phi_1(\xi), ..., \phi_N(\xi))$, for some N.

(i, ii) are established in Appendix 7.3.2. (iii) follows immediately from them and the implicit function theorem (IFT in the sequel).

As usual, we call regular an interior SSE such that $D_\phi \Phi(\phi; \overline{\xi})$ has full rank (implicitly, this requires that (i) above holds, otherwise, $\Phi(\phi; \overline{\xi})$ may be non differentiable). If each interior SSE of an economy is regular, we call the economy regular.

4. Efficiency properties of interior SSE allocations

We start with an informal discussion. As common in the literature, we define efficiency in terms of steady state total, discounted, expected surplus. More problematic is to select the appropriate set of policy instruments. An intuitively appealing notion of constrained optimality (CO) requires that the SSE allocation cannot be improved upon just by changing the set of people getting educated and the "market tightness" variables ϕ (i.e., unemployment and vacancies). Unfortunately, such a notion would be useless in our context and for our purposes. An SSE allocation, to be CO, must be an optimal solution to the planner's problem, given the specific selection of the set Θ^{0e}. In the canonical one-sector, random matching model, SSE are typically constrained inefficient, unless the "market power" weight β happens to satisfy the Hosios (1990) condition: $\beta = |\eta_{q^k}(\phi)|$, where $\eta_{q^k}(\phi)$ is the elasticity of $q^k(\phi)$ with respect to ϕ^k. Given Θ^{0e}, our model reduces to a pair of disjoint random matching economies. Therefore, a necessary condition for a SSE to be CO is $\beta = |\eta_{q^k}(\phi)|$, for each k. Hence, SSE allocations typically are not CO, as long as β is an exogenous parameter. This obscures the nature of inefficiencies specifically related to the educational choices of the agents, if any. To avoid this problem, we propose a different efficiency concept, Weak Constrained Optimality (WCO). With our definition, the planner can choose any measurable subset Θ^{0e}. This choice induces a pair ϕ, obtained as SSE of the two (now, separated) sub-economies e and ne.[14] In a one-sector model, interior SSE are trivially WCO, because they are globally unique and the constraint set of the planner reduces to a single point, the SSE itself. Hence, the notion of WCO is extremely weak, and, consequently, WCO allocations do not have a strong appeal from the normative viewpoint. This criterion, however, is useful. First, it allows us to pinpoint sources of inefficiency just

[14] Our notion of WCO is somewhat related to the one commonly used in the literature on general equilibrium with incomplete markets (see, Geanakoplos and Polemarchakis (1986)). In both cases, the planner chooses the investment portfolios, taking into account the induced adjustment of the endogenous equilibrium variables (prices there, ϕ here). We are completely agnostic about the (far from trivial) problem of the existence of WCO allocations, that is not really germane to the issue under consideration.

related to the two-sector structure of the economy, and to the private nature of the investments in education. Secondly, the weaker the notion of CO, the stronger the inefficiency result.

We restrict the analysis to steady states, and assume $r' = 0$, so that $r = \gamma$ (or, rather, we consider the limit case for r' converging to zero). This entails a loss of generality, but it allows us to sidestep issues related to dynamic optimality versus optimality of the steady states, and it is standard in this literature.

In the sequel, the length of schooling, T, is fixed, hence omitted from notation.

The planner's objective function, $P'(u, o, \Theta^{0e}; \xi)$, is the steady state, discounted, expected total surplus, net of vacancy costs and of the direct costs of education. Its explicit form is given by (12), in Appendix 7.4. The planner's policy instruments are the choice of a measurable subset of Θ^{0e} and of the pair (u, o). The planner faces three constraints:

$a.$ $\quad u^e = \frac{\gamma}{\gamma + \pi^e(\phi)}$;

$b.$ $\quad u^{ne} = \frac{\gamma}{\gamma + \pi^{ne}(\phi)}$;

$c.$ $\quad \Phi_{\Theta^{0e}}(\phi; \xi) = 0.$

The last constraint differs from the equilibrium condition $\Phi(\phi; \xi) = 0$. In $\Phi_{\Theta^{0e}}(\phi; \xi)$, Θ^{0e} is selected by the planner, while, in $\Phi(\phi; \xi)$, it is implicitly given by the additional condition $G(\theta, \phi; \xi) \geq 0$, for each $\theta \in \Theta^{0e}$. Given the constraints $(a - b)$, the policy instruments actually reduce to Θ^{0e} and ϕ. Also, we are implicitly imposing symmetry in the treatment of agents of the same type θ, because it always holds, at a WCO allocation.

For completeness, we reformulate in our set-up the standard notion of CO and the inefficiency result already mentioned.

Define as $P(\phi, \Theta^{0e}; \xi)$ the objective function obtained replacing into $P'(u, o, \Theta^{0e}; \xi)$ the stationary values of (u, o), given by $(a - b)$ above and by $o^k = \phi^k u^k$, for each k.

Definition 2. A pair $(\overline{\phi}, \overline{\Theta}^{0e})$ is Constrained Optimal (CO) if and only if it is solution to the problem

$$choose\ (\phi, \Theta^{0e}) \in \arg\max P(\phi, \Theta^{0e}; \xi).$$

Proposition 1. Under A1-A6, there is an open, dense subset $\Xi' \subset \Xi_T^{reg}$ such that, for each $\xi \in \Xi'$, each interior SSE is not CO.

The proof of Prop. 1 is straightforward (given Theorem 2 above) and, therefore, omitted.

We obtain the notion of WCO by introducing into the planner's optimization problem the additional constraint (c), defined above.

Definition 3. A pair $(\overline{\phi}, \overline{\Theta}^{0e})$ is Weakly Constrained Optimal (WCO) if and only if it is a solution to the problem

$$choose\ (\phi, \Theta^{0e}) \in \arg\max P(\phi, \Theta^{0e}; \xi)\ subject\ to\ \Phi_{\Theta^{0e}}(\phi; \xi) = 0.$$

In Appendix 7.4 we establish our key result.

Theorem 3. Under A1-A6, there is an open, dense subset of economies, $\overline{\Xi} \subset \Xi_T^{reg}$, such that, for each $\xi \in \overline{\Xi}$, each interior SSE is not WCO.

Remark 2. Throughout the paper, β is an exogenous parameters. As we will see later on, the value of $(\beta + \eta_{q^k}(\phi))$ plays a role in determining the lack of WCO of SSE and, most important, the nature of the inefficiency. However, the Hosios′ condition is neither necessary, nor sufficient, to guarantee WCO.

Remark 3. Thm. 3 holds generically in the set of regular economies with interior SSE. We have formally established existence of a (possibly) small subset of economies with these properties. However, this last theorem does not rest in any substantive sense on Thm. 1. Its result holds generically in the set of regular economies with interior SSE. Moreover, its proof rests heavily on differentiability. This property is never at issue here. In the planner's problem what matters are the derivatives $\frac{\partial \phi^k}{\partial \theta_m}$, for $k = e, ne$, obtained by the IFT applied to the constraint $\Phi_{\Theta^{0e}}(\phi; \xi) = 0$, at the values θ_m selected by the planner. While it is possible that $rank\ D_\phi \Phi(\phi; \xi) < 2$, $rank D_\phi \Phi_{\Theta^{0e}}(\phi; \xi) = 2$, always. Therefore, $\frac{\partial \phi^k}{\partial \theta}|_{\theta=\theta_m}$, for $k = e, ne$, are always well-defined, here.

The literature has identified three different possible sources of constrained inefficiency. First, when $(\beta + \eta_{q^k}(\phi)) \neq 0$ at the SSE, the equilibrium is constrained inefficient, because agents do not internalize the congestion externality. Secondly, with investments in human capital, there may be an "hold up" effect, stressed by Acemoglu (1996) in a different, but related model: educated workers do not receive the full return on their investment, because of the wage setting mechanism, and of the irreversibility of their investment. In his model, this always induces underinvestment in education.[15] Finally, when a subset of workers switches from one sector to the other, the expected productivities in both sectors may vary, and with them the creation of vacancies. This is an additional potential source of inefficiency and, following Charlot and Decreuse (2005), we call it composition effect.[16]

Our notion of WCO neutralizes sector-by-sector congestion externalities. The other two kinds of sources of inefficiency are potentially active.

Let's make precise our notions of over and undereducation. As in the proof of Thm. 3, and without any loss of generality, we restrict the planner to choose sets Θ^{0e} given by the union of a finite collection of intervals $[\theta_m, \theta_{m+1}]$. Replace into its objective function the pair of C^1 functions $\phi^k(\theta_1, ..., \theta_m)$, $k = ne, e$, obtained applying the IFT to the constraint $\Phi_{\Theta^{0e}}(\phi; \xi) = 0$. As mentioned in Remark 3 above, this pair of functions always exists. Then, the modified planner's optimization problem is

$$\max_{[\theta_1, ..., \theta_m, ..., \theta_M]} \overline{P}(\theta_1, ..., \theta_M; \xi) \equiv P(\phi^e(.), \phi^{ne}(.), (\theta_1, ..., \theta_M); \xi). \tag{10}$$

Define $\chi(\theta) = 1$, if $\theta_m \in [\theta_m, \theta_{m+1}] \subset \Theta^{0e}$, $\chi(\theta) = 2$, if $\theta_m \in [\theta_{m-1}, \theta_m] \subset \Theta^{0e}$.

Definition 4. A SSE of the economy $\xi \in \Xi$ exhibits (local) undereducation (overeducation) at $\theta_m \in G_\phi^{-1}(0)$ if and only if $(-1)^{\chi(\theta_m)} \frac{\partial \overline{P}}{\partial \theta_m} > 0$ (if and only if $(-1)^{\chi(\theta_m)} \frac{\partial \overline{P}}{\partial \theta_m} < 0$).

This formulation will become handy in the sequel and the basic idea is really straightforward. There is overeducation if (locally) we increase the total surplus by shrinking the set of agents investing in education. If θ_m is the lower bound of an interval $[\theta_m, \theta_{m+1}] \subset \Theta^{0e}$, this requires $\frac{\partial \overline{P}}{\partial \theta_m} > 0$. If θ_m is an upper bound, it requires $\frac{\partial \overline{P}}{\partial \theta_m} < 0$, i.e., it requires $(-1)^{\chi(\theta_m)} \frac{\partial \overline{P}}{\partial \theta_m} < 0$.

[15] Acemoglu's economy is always characterized by full employment, so that there are no congestion externalities at play. Constrained inefficiency is due to lack of contractibility of investments and the hold-up problem. His results are relevant in our set-up, because a somewhat similar "hold-up" effect is at play here.

[16] In their model, it always induces overinvestment in education, because of the efficiency criterion that they adopt.

The (necessary) first order conditions of the modified planner's optimization problem (10) are

$$\frac{\partial \overline{P}}{\partial \theta_m} = \frac{\partial P}{\partial \theta_m} + \left(\frac{\partial P}{\partial \phi^e} \frac{\partial \phi^e}{\partial \theta_m} + \frac{\partial P}{\partial \phi^{ne}} \frac{\partial \phi^{ne}}{\partial \theta_m} \right) = 0, \text{ for each } \theta_m.$$

Thus, $\frac{\partial \overline{P}}{\partial \theta_m}$ is the sum of two terms, capturing the direct and indirect effects of changes in θ_m on the objective function. As in Appendix 7.4, define the map

$$\begin{aligned} T(\rho, \theta, \phi; \xi) &= \alpha(\theta) \frac{\rho \pi^e(\phi) f^e(\theta) + \gamma b^e(\theta)}{\gamma + \rho \pi^e(\phi)} + \left(1 - e^{\gamma T}\right) c(\theta) \\ &\quad + \left(1 - e^{\gamma T} - \alpha(\theta)\right) \frac{\rho \pi^{ne}(\phi^{ne}) f^{ne}(\theta) + \gamma b^{ne}(\theta)}{\gamma + \rho \pi^{ne}(\phi)}, \end{aligned}$$

where, at $\rho = \beta$, $T(\beta, \theta, \phi; \xi) = G(\theta, \phi; \xi)$, while, at $\rho = 1$, $T(1, \theta, \phi; \xi)$ is the social gain (net of the total cost of education) of the investment in education of agent θ, i.e., the relevant variable from the planner's viewpoint.

Evidently, if an interior SSE is WCO, it must satisfy the necessary first order conditions of optimization problem (10). Then, to understand the nature of the inefficiency, it suffices to study the sign of $\frac{\partial \overline{P}}{\partial \theta_m}$ at each $\theta_m \in G_\phi^{-1}(0)$, evaluated at an interior SSE. By direct computation, using the expressions for $\overline{P}(\theta_1, ..., \theta_M; \xi)$ which is obtained from eq. (12) (in Appendix 7.4), at each interior SSE,

$$(-1)^{\chi(\theta m)} e^{\gamma T} \frac{\partial P}{\partial \theta_m} = T(1, \theta_m, \phi; \xi) - \left(\frac{\alpha(\theta_m) \gamma v^e \phi^e}{\gamma + \pi^e(\phi)} + \frac{\left(1 - e^{\gamma T} - \alpha(\theta_m)\right) \gamma v^{ne} \phi^{ne}}{\gamma + \pi^{ne}(\phi)} \right).$$

The first term, $T(1, \theta_m, \phi; \xi)$, is related to the "hold up" problem stressed in Acemoglu (1996), because, when $\beta = 1$, the value of $T(1, \theta_m, \phi; \xi)$ coincides with the one of $G(\theta_m, \phi; \xi)$, and, therefore, is zero. Assume $b^k(\theta) = 0$. Then, $T(1, \theta_m, \phi; \xi)$ is positive if $\pi^e(\phi)$ is not "too large" compared to $\pi^{ne}(\phi)$.[17]

The second term is the difference between the discounted expected vacancy costs in the two markets. It is obviously nil if there is a unique labor market and education does not require time. If this term is sufficiently small, and $(\pi^e(\phi) - \pi^{ne}(\phi))$ not too large, $(-1)^{\chi(\theta)} \frac{\partial P}{\partial \theta_m} > 0$, so that the direct effect induces undereducation. On the other hand, if $(\pi^e(\phi) - \pi^{ne}(\phi))$ is positive and sufficiently large, we may have $(-1)^{\chi(\theta)} \frac{\partial P}{\partial \theta_m} < 0$, i.e., overeducation at θ_m. Bear in mind that, by definition of WCO, the direct effect does not depend in any way upon $\left(\beta + \eta_{q^k}(\phi)\right)$, because it is computed at the given SSE vector ϕ^*. Hence, it can be different from zero even if the Hosios' condition holds, for each k.

The second, indirect, component captures the welfare impact of the composition effect. By direct computation, rearranging terms, and using the fact that, at a SSE, $\Phi(\phi; \xi) = 0$, we obtain that, at each θ_m,

$$(-1)^{\chi(\theta_m)} \frac{\partial P}{\partial \phi^k} \frac{\partial \phi^k}{\partial \theta_m} = \left[\frac{\gamma v^k \mu \left(\Theta^{k\alpha}(\phi; \xi)\right) \left(\beta + \eta_{q^k}(\phi)\right)}{(1 - \beta)\left(\gamma + \pi^k(\phi)\right)} \right] \frac{\partial \phi^k}{\partial \theta_m}, \text{ each } k,$$

[17] At each ρ such that $T(\rho, \theta, \phi; \xi) = 0$,

$$\frac{\partial T}{\partial \rho} = \frac{\gamma \left[\frac{\alpha(\theta_m) \rho \pi^e(\phi) f^e(\theta_m)}{\gamma + \rho \pi^e(\phi)} + \left(1 - e^{\gamma T} - \alpha(\theta_m)\right) \frac{\pi^{ne}(\phi) \rho f^{ne}(\theta_m)}{\gamma + \rho \pi^{ne}(\phi)} \frac{\gamma + \rho \pi^e(\phi)}{\gamma + \rho \pi^{ne}(\phi)} \right]}{\rho \left(\gamma + \rho \pi^e(\phi)\right)} > 0$$

for $\pi^e(\phi)$ sufficiently close to (or smaller than) $\pi^{ne}(\phi)$. Given that $T(\rho = \beta, \theta, \phi; \xi) = 0$, this implies $T(1, \theta, \phi; \xi) > 0$.

where, for each k, $\frac{\partial \phi^k}{\partial \theta_m}$ is computed applying the IFT to the map $\Phi_{\Theta^{0e}}(\phi;\xi)$, the relevant one from the planner's viewpoint. For each k, this term is nil if and only if either Hosios' condition holds or $\frac{\partial \phi^k}{\partial \theta_m} = 0$. The Hosios' condition comes back into play because of changes of the equilibrium value of ϕ due to changes in θ, even if our notion of efficiency is constructed to neutralize the canonical (i.e., given $\Theta^{k\alpha}(\phi;\xi)$, each k) Hosios' effect. Assume $(\beta + \eta_{q^k}(\phi)) > 0$, each k, so that unemployment is above its constrained optimum (according to the usual criterion) in both sectors.[18] Then, a change in θ_m which decreases unemployment in each sector has a positive effect on surplus. By the IFT, and direct computation,

$$\begin{aligned}
\frac{\partial \phi}{\partial \theta_m} &= -\left[\frac{\partial \Phi_{\Theta^{0e}}}{\partial \phi}\right]^{-1}\left[\frac{\partial \Phi_{\Theta^{0e}}}{\partial \theta_m}\right] \\
&= (-1)^{\chi(\theta_m)} e^{-\gamma T}\begin{bmatrix} \alpha(\theta_m)\frac{(f^e(\theta_m)-b^e(\theta_m))-F^e(\phi;\xi)}{\mu(\Theta^{e\alpha}(\phi;\xi))}\left(-\frac{\partial \Phi^e_{\Theta^{0e}}}{\partial \phi^e}\right)^{-1} \\ [1-e^{\gamma T}-\alpha(\theta_m)]\frac{(f^{ne}(\theta_m)-b^{ne}(\theta_m))-F^{ne}(\phi;\xi)}{\mu(\Theta^{ne\alpha}(\phi;\xi))}\left(-\frac{\partial \Phi^{ne}_{\Theta^{0e}}}{\partial \phi^{ne}}\right)^{-1}\end{bmatrix}.
\end{aligned}$$

It is easy to check that $\frac{\partial \Phi^k_{\Theta^{0e}}}{\partial \phi^k} < 0$, each k. Given that $\left(1-e^{\gamma T}-\alpha(\theta_m)\right) < 0$,

$$sign\begin{bmatrix}\frac{\partial \phi^e}{\partial \theta_m} \\ \frac{\partial \phi^{ne}}{\partial \theta_m}\end{bmatrix} = (-1)^{\chi(\theta_m)} sign\begin{bmatrix}[(f^e(\theta_m)-b^e(\theta_m))-F^e(\phi;\xi)] \\ -[(f^{ne}(\theta_m)-b^{ne}(\theta_m))-F^{ne}(\phi;\xi)]\end{bmatrix}.$$

Hence, the sign of the indirect effect depends, for each k, upon the signs of $(\beta + \eta_{q^k}(\phi))$, and of $\frac{\partial \phi}{\partial \theta_m}$. The second determines the effect on the rate of arrival of matches (hence, on unemployment) of a marginal change in the set of agents investing in education. The first determines the welfare effect of an increase in the value of ϕ.

Generally speaking, it is hard to discriminate between overeducation and undereducation, also because, when there are several $\theta_m \in G_\phi^{-1}(0)$, the SSE may be characterized by overeducation at some of them, by undereducation at some others. The case where there is always a unique SSE threshold value $\overline{\theta}$ defining $\Theta^{0e}(\phi;\xi)$ is considered in the next section. The results of this section, however, add to the previous literature for at least three reasons. First, they clarify the role of the Hosios' condition. The classical Hosios' (1990) effect is absent, because of our definition of WCO. However, the same condition comes back into play because of the composition effect. The hold-up problem discussed in Acemoglu (1996) plays a role in our set-up,[19] too. However, it does not guarantees that equilibria are characterized by undereducation, hence, that there is always a positive externality related to investments in human capital.

5. Two polar cases: Ability and education as complements and substitutes

To conclude, we consider the two polar cases where, at each interior SSE, there is always a unique threshold $\overline{\theta} \in G_\phi^{-1}(0)$. T is fixed throughout the section. We assume $(\frac{\partial f^e}{\partial \theta}, \frac{\partial f^{ne}}{\partial \theta}, \frac{\partial \alpha}{\partial \theta}) >> 0$ and, therefore, we can now interpret θ as some measure of ability in all the relevant activities of the workers. We define as complementarity the case where only the high θ types invest in education.

[18] In the one-sector model, the term in square brackets is the FOC of the planner's optimization problem. This is where the Hosios condition comes from.

[19] In his model, both human capital (in terms of efficiency units) and physical capital are elastically supplied. In a two-sector version of the same model, the welfare properties of the model are sharply different, because of self-selection in the two sectors and of the composition effect (see, Mendolicchio, Paolini, Pietra (2009)).

As substitutability, the one where just low θ agents invest. A priori, both cases are plausible, because what matters is how the comparative advantages in the two jobs vary with θ.

Definition 5. An economy $\xi \in \Xi$, is characterized by complementarity (substitutability) between ability and education if and only if, at each interior SSE $\phi(\xi)$, $\Theta^{0e}(\phi(\xi);\xi) = \left[\overline{\theta}(\phi;\xi), \theta_h\right]$ $(\Theta^{0e}(\phi(\xi);\xi) = \left[\theta_\ell, \overline{\theta}(\phi;\xi)\right])$.

For instance, the economies considered in Charlot and Decreuse (2005) are characterized by complementarity.

We start providing some restrictions on the fundamentals which allow for a (partial) characterization of complementarity vs. substitutability. Evidently, a sufficient condition for uniquess of the threshold value $\overline{\theta}$ is that the sign of $\frac{\partial G}{\partial \theta}$ is invariant at each $\theta \in G_\phi^{-1}(0)$, for each possible SSE ϕ. Thus, we obtain complementarity if $\frac{\partial G}{\partial \theta}$ is always positive, substitutability if it is always negative. Proposition 2 provides a sufficient condition for the invariance of its sign at each SSE.

First, we introduce a set of additional, simplifying, assumptions:

Assumption 7 (A7): $b^k(\theta) = c(\theta) = 0$, each k and θ. Moreover, $\nu^e = \nu^{ne}$ and $q^e(.) = q^{ne}(.)$,[20] $(\frac{\partial f^e}{\partial \theta}, \frac{\partial f^{ne}}{\partial \theta}, \frac{\partial \alpha}{\partial \theta}) >> 0$.

Also, let's define as $\eta_\alpha(\theta)$, and $\eta_{f^k}(\theta)$ the elasticities (with respect to θ) of the functions $\alpha(\theta)$ and $f^k(\theta)$, for each k.

Proposition 2. Under A1-A7, we obtain

a. complementarity: for each $\xi \in \Xi$ such that $\eta_\alpha(\theta) > 0$ and $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)} \geq 1$, for each θ;

b. substitutability: for each $\xi \in \Xi$ such that $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)} < 1$, and $\eta_\alpha(\theta)$ is sufficiently small for each θ.

All the proof of this section are in Appendix 7.5.

The discussion of WCO for the general case has already pointed out the role of the sign of $(\beta + \eta_{q^k}(\phi))$ in affecting the nature of inefficiency. For a general matching function, this sign depends upon the SSE value of ϕ. However, in many applications, the analysis is restricted to constant elasticity functions (see Petrongolo and Pissarides (2001)). Hence, it is meaningful to focus on this case.[21]

As above, we study the welfare properties of SSE assuming that $r\prime = 0$, i.e., $r = \gamma$.

Proposition 3. Assume that complementarity holds, and that $\alpha(\theta)$ and $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ are sufficiently large for each θ. Then, each SSE is characterized by overeducation if:

a. $\beta \geq |\eta_{q^k}|$;

b. $\beta < |\eta_{q^k}|$ and $\eta_{f^e}(\theta)$ is sufficiently small for each θ.

[20] With $q^e(.) = q^{ne}(.)$, we mean that the matching functions are identical in the two sectors. Evidently, as long as $\phi^e \neq \phi^{ne}$, $q^e(\phi) \neq q^{ne}(\phi)$

[21] According to Petrongolo and Pissarides (2001, p. 393), the range of the most plausible values of $\eta_{q^k}(\phi)$ is $(-0.5, -0.3)$. The value of β has been estimated for several countries. Most of the results suggest that it is fairly small (see Yashiv (2003, 2006), Cahuc, Postel-Vinay, and Robin (2006) and other references quoted therein). However, Cahuc, Postel-Vinay, and Robin (2006) report values of β^{ne} around 0.1, but substantially larger values for β^e. Flinn and Mabli (2008) reports relatively high values of β. Hence, it is worthwhile to consider both $\beta \geq |\eta_{q^k}|$ and $\beta < |\eta_{q^k}|$

With complementarity, the direct effect on welfare of an increase in the value of the threshold $\overline{\theta}$ is always positive, under the assumptions of Prop. 3. This guarantees that inefficiency is due to overeducation if $\beta \geq |\eta_{q^k}|$, because the indirect effect is also positive: an increase in $\overline{\theta}$ induces an increase in the equilibrium value of $\pi^k(\phi)$ in both markets. When $\beta \geq |\eta_{q^k}|$, this is welfare improving. On the other hand, if $\beta < |\eta_{q^k}|$ the composition effect has a negative impact on welfare. A sufficiently small value of $\eta_{f^e}(\theta)$ guarantees that the direct effect dominates.

Let's now turn to the case of substitutability.

Proposition 4. Assume that substitutability holds, and that $\alpha(\theta)$ is sufficiently large and $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ is not "too large", for each θ. Then each SSE is characterized by undereducation if:
a. $\beta \geq |\eta_{q^k}|$;
b. $\beta < |\eta_{q^k}|$ and $\eta_{f^e}(\theta)$ and $\eta_{f^{ne}}(\theta)$ are sufficiently small, for each θ.

We have studied efficiency with reference to the optimal solution of an abstract optimization problem, where the planner selects the subset of agents investing in education. A natural question concerns the possible implementation of welfare improvements obtained using standard policy instruments. In our set-up, the simpler possible instrument are taxes and subsidies on the direct costs of education. We conclude showing that sufficiently small taxes (or subsidies) on the direct costs of education are welfare improving and can mitigate the adverse effects of over/undereducation. This requires us to evaluate the impact on welfare of the change in the SSE variables induced by a change in the exogenous policy instrument, hence, as a preliminary step, to determine the comparative statics properties of SSE.

Given that the threshold $\overline{\theta}$ is unique, we can describe interior SSE by the system of eqs.

$$\Psi\left(\overline{\theta},\phi;\xi\right) \equiv \left(\Phi^e\left(\overline{\theta},\phi;\xi\right),\Phi^{ne}\left(\overline{\theta},\phi;\xi\right),G\left(\overline{\theta},\phi;\xi\right)\right)=0.$$

Let $\frac{\partial\overline{\theta}_G}{\partial\phi^k}$, for each k, and $\frac{\partial\overline{\theta}_G}{\partial\xi}$ define the derivatives of the map $\overline{\theta}(\phi;\xi)$ obtained, at the SSE, considering only the map $G(\theta,\phi;\xi)=0$. Let $\frac{\partial\overline{\theta}}{\partial\xi}$ denote instead the derivative of the actual equilibrium map $\overline{\theta}(\xi)$. By the IFT, $\left[\frac{\partial\phi^e}{\partial\xi},\frac{\partial\phi^{ne}}{\partial\xi},\frac{\partial\overline{\theta}}{\partial\xi}\right]^T=-D_{(\phi,\theta)}\Psi^{-1}D_\xi\Psi$, which can be written as

$$\det D_{\left(\phi,\overline{\theta}\right)}\Psi\left(\phi;\xi\right)=\frac{\partial G}{\partial\overline{\theta}}\left[\frac{\partial\Phi^{ne}}{\partial\phi^{ne}}\frac{\partial\Phi^e}{\partial\phi^e}+\frac{\partial\Phi^{ne}}{\partial\phi^{ne}}\frac{\partial\Phi^e}{\partial\overline{\theta}}\frac{\partial\overline{\theta}_G}{\partial\phi^e}+\frac{\partial\Phi^e}{\partial\phi^e}\frac{\partial\Phi^{ne}}{\partial\overline{\theta}}\frac{\partial\overline{\theta}_G}{\partial\phi^{ne}}\right]. \tag{11}$$

We will further assume that, with complementarity $\left[\frac{\partial\Phi^{ne}}{\partial\theta}\frac{\partial\overline{\theta}}{\partial\phi^{ne}}+\frac{\partial\Phi^{ne}}{\partial\phi^{ne}}\right]<0$, i.e., that the total effect of an increase in ϕ^k on expected profits in sector k (taking into account the change in the composition of the pool of workers) is negative. Similarly, with substitutability, we impose $\left[\frac{\partial\Phi^e}{\partial\theta}\frac{\partial\overline{\theta}}{\partial\phi^e}+\frac{\partial\Phi^e}{\partial\phi^e}\right]<0$. It is easy to check that this (together with $\alpha(\theta)$ large) implies that $\det D_{\left(\phi,\overline{\theta}\right)}\Psi(\phi;\xi)$ is positive with complementarity, negative with substitutability. Also, one can check that this additional restrictions are compatible with the other maintained assumptions.

We define the shocks to technologies, direct costs of education, probability of graduation and matching functions in terms of a one-dimensional parameterization, defined, for instance, by $\alpha(\theta,a)=(1+a)\alpha(\theta)$. Shocks to vacancy costs are defined in the obvious way. For the function $c(\theta)$, we focus on the case of θ-invariant, additive shocks.[22] The proofs of Prop. 5 and 6 follow by a strainghtforward (but very tedious and, hence, omitted) computation.

[22] Given that $c(\theta)=0$, for each θ, multiplicative shocks are meaningless. We omit shocks to (b^e,b^{ne}) because they do not have a clear effect on the equilibrium variables.

Proposition 5. Under A1-A7, if $\left(\frac{\partial \Phi^{ne}}{\partial \phi^{ne}} + \frac{\partial \Phi^{ne}}{\partial \overline{\theta}} \frac{\partial \overline{\theta}_G}{\partial \phi^{ne}}\right) < 0$, and $\alpha(\theta)$ is sufficiently large, for each θ, in the case of complementarity, the following sign restrictions hold

$$\begin{bmatrix} \diagdown \xi = & f^e & f^{ne} & c & \alpha & v^e & v^{ne} & q^e & q^{ne} \\ \frac{\partial \phi^e}{\partial \xi} & ? & + & + & - & - & - & ? & + \\ \frac{\partial \phi^{ne}}{\partial \xi} & - & + & + & - & + & - & - & + \\ \frac{\partial \overline{\theta}}{\partial \xi} & - & + & + & - & + & - & - & + \end{bmatrix}$$

A sufficient condition for $\frac{\partial \phi^e}{\partial f^e} < 0$ is $\left(\frac{\partial \Phi^e}{\partial f^e} + \frac{\partial \Phi^e}{\partial \overline{\theta}} \frac{\partial \overline{\theta}}{\partial f^e}\right) < 0$. A sufficient condition for $\frac{\partial \phi^{ne}}{\partial q^e} < 0$ is $\left(\frac{\partial \Phi^e}{\partial q^e} + \frac{\partial \Phi^e}{\partial \overline{\theta}} \frac{\partial \overline{\theta}_G}{\partial q^{ne}}\right) < 0$.

The results above can be interpreted in terms of the Charlot and Decreuse's (2005) composition effect. Changes in the exogenous parameters making, ceteris paribus, the market for uneducated workers more attractive (i.e., $df^{ne} > 0, dc > 0, d\alpha < 0, dv^{ne} < 0, dq^{ne} > 0$) always increase both ϕ^e and ϕ^{ne}. This is because they attract (comparatively) higher ability individuals to this market, improving the (conditional) expected product in both sectors. On the other hand, consider, for instance, the effect of a positive technological shock in sector e. It induces an increase in the demand for education. Necessarily, this raises unemployment in sector ne. Its effect on unemployment in sector e is ambiguous. The direct effect is positive, while the composition effect is negative. If this is sufficiently large (so that $\left(\frac{\partial \Phi^e}{\partial f^e} + \frac{\partial \Phi^e}{\partial \overline{\theta}} \frac{\partial \overline{\theta}_G}{\partial f^e}\right)$ is negative), we can actually have an increase in the equilibrium unemployment level. A similar argument holds for changes in $q^e(\phi)$.

The case of substitutability can be discussed in a similar way, once one takes into account that, now, an increase in the investment in education improves the average productivity in both sectors.

Proposition 6. Under A1-A7, if $\left(\frac{\partial \Phi^e}{\partial \phi^e} + \frac{\partial \Phi^e}{\partial \overline{\theta}} \frac{\partial \overline{\theta}_G}{\partial \phi^e}\right) < 0$, $\frac{\partial \alpha(\theta)}{\partial \theta} > 0$, and $\alpha(\theta)$ is sufficiently large, for each θ, in the case of substitutability, the following sign restrictions hold

$$\begin{bmatrix} \diagdown \xi = & f^e & f^{ne} & c & \alpha & v^e & v^{ne} & q^e & q^{ne} \\ \frac{\partial \phi^e}{\partial \xi} & + & - & - & ? & - & + & + & - \\ \frac{\partial \phi^{ne}}{\partial \xi} & + & ? & - & + & - & - & + & ? \\ \frac{\partial \overline{\theta}}{\partial \xi} & + & - & - & ? & - & + & + & - \end{bmatrix}.$$

A sufficient condition for $\frac{\partial \phi^{ne}}{\partial \xi} < 0$ is $\left(\frac{\partial \Phi^{ne}}{\partial f^{ne}} + \frac{\partial \Phi^{ne}}{\partial \theta} \frac{\partial \overline{\theta}_G}{\partial f^{ne}}\right) < 0$.

Putting together the comparative statics results of the last two Proposition and the welfare results of Prop. 3 and 4, we can identify the natural policy tool to implement welfare improvements. Consider the case of complementarity with overeducation. The introduction of a tax Δc on the direct cost of education (whose revenues are redistributed as lump-sum taxes) increases the SSE threshold value $\overline{\theta}$ and the pair (ϕ^e, ϕ^{ne}), by Prop. 5. It is straightforward to show that, if overeducation holds, this policy is welfare improving. Similarly for the case of substitutability and undereducation. To summarize,

Proposition 7. Under the assumptions of Prop. 2-6, a tax on the direct cost of education (whose revenues are reallocated with lump-sum taxes) is welfare improving in the case of complementarity, welfare reducing in the one of substitutability.

6. Conclusion

We have embedded the Roy's model in a random matching environment. From a generic viewpoint, the resulting economy is well-defined (i.e., there is a SSE, under some restrictions on the - exogenous - length of the education process). Interior SSE are generically inefficient according to a very weak notion of constrained efficiency. Given the technique of proof adopted, these properties are robust to many possible extensions of the model. The more stylized, but still fairly general, versions of the model considered in the last section allow us to obtain reasonably sharp comparative static properties of SSE, and a partial characterization of inefficiency in terms of overeducation or undereducation. The nexus between comparative statics properties and the nature of inefficiency makes the model potentially testable and allows (under the proper set of restrictions) to identify the policy instrument sufficient to implement welfare improvements.

A key feature of the model is the role of the composition effect. As pointed out in Charlot and Decreuse (2005), investments in education allows agents to self-select themselves into one of the labor markets. In economies with frictions, generically, this has relevant consequences, which are ruled out by assumption in economies where investments in education translate into an increase in the number of efficiency units of the labor endowments.

An essential ingredient of our model is the assumption that matching is at random. The extension of the analysis to economies with directed search is left for future research.

7. Appendices

7.1. Stochastic structure

Assume that, at each t, a population is described by an interval $\Theta(t)$ endowed with the Lebesgue measure, and that, for each $i \in \Theta(t)$, "death" follows a Poisson process with arrival rate γ, and i.i.d. realizations. Then, it is often argued that this implies $\frac{\partial \mu(\Theta(t))}{\partial t} = -\gamma \mu\left(\Theta\left(t\right)\right)$. This statement is not correct, see Judd (1985), and Feldman and Gilles (1985). More recent contributions on this issue include Al-Najjar (2004), and Alòs-Ferrer (1999, 2002). Duffie and Sun (2007), Sun (2007), Sun and Zhang (2009) provide a construction which guarantees the validity of a modified form of LLN. This requires to replace the standard unit interval-Lebesgue measure framework, and to modify the independence assumption.

If one wants to preserve as index set the interval $[0,1]$ endowed with the Lebesgue measure, the i.i.d. assumption is the core of the problem. In many economic applications, we just need to have, at the same time, individual uncertainty and aggregate certainty. The easiest way to obtain this property in the usual set-up is to drop the i.i.d. assumption. This approach has several advantages: First, it allows us to maintain the same structure of the random matching literature. Second, it avoids technical complexities. Third, it can be easily applied to Poisson processes. Following Allen (1985, p. 96-97) and Feldman and Gilles (1985, p. 28-29), given $x \in [0,1]$, and a realization ω of a r.v. $\widetilde{\omega}$ uniformly distributed on $[0,1]$, define the interval

$$[\omega \oplus x] \equiv [\max\{0, \omega - x\}, \omega] \cup [1 - \min\{0, \omega - x\}, 1].$$

Termination and creation of job matches are described by Poisson processes with arrival rates γ (exogenous), $\pi^k(\phi^k)$ and $q^k(\phi^k)$ (endogenous). The distinction between endogenous and exogenous arrival rates is irrelevant, here, at least as long as we consider steady states. We focus

on the termination of job matches, and start with discrete time. Consider a discrete process with time intervals of length Δ. Let $\widehat{\theta}(t)$ be the set of agents of type θ alive at the beginning of the period. Its measure is $\mu(\widehat{\theta}(t)) = 1$. In each period, a r.v. $\widetilde{\omega} \in [0,1]$ realizes. If, at time interval $[t, t+\Delta,]$ $\widetilde{\omega}_t = \omega_t$, death occurs for each $i \in \omega_t \oplus \Delta\gamma$. Hence, for each θ, and each period, a fraction $\Delta\gamma$ of agents dies, and, for each θ_i, the probability of death in (if alive at the beginning of the period) is $\Delta\gamma$. Consider now the limit for $\Delta \to 0$. By definition of derivative, for each $\theta \in \Theta^0$, the rate of change of the population due to "death" is $-\gamma$, so that A1-(b) holds. Moreover, by a standard argument, for each θ_i, "death" occurs according to a Poisson process with arrival rate γ, so that A1-(a) holds.

Consider now the r.v. "graduation". Pick $\theta \in \Theta^{0e}$ and, as above, consider a discrete time process with intervals of length Δ. A set of measure $\gamma\Delta$ of agents of type θ is born in the interval $[t, t+\Delta]$. To fix ideas, assume that, in that time interval, the realization ω satisfies $\omega > \gamma\Delta$, so that the interval of agents born in the period is $[\omega - \gamma\Delta, \omega]$. Let $[\underline{n}, \overline{n}]$ index the set of agents born in $[t, t+\Delta]$ and still alive after T time intervals. Consider now period $[t+T\Delta, t+(T+1)\Delta]$ and assume that the value $\widehat{\omega}$ of the r.v. $\widetilde{\omega}$ realizes. Define the interval $\widehat{\omega}[\overline{n} - \underline{n}] \oplus \alpha(\theta)[\overline{n} - \underline{n}]$ and assume that agents in this interval graduate. Evidently, each agents graduates with probability $\alpha(\theta)$ and a fraction $\alpha(\theta)$ of the set of agents born at $[t, t+\Delta]$ and alive at $[t+T\Delta, t+(T+1)\Delta]$ graduates. Taking the limits for Δ converging to 0, we obtain A2.

7.2. Transversality theorem

Several of the proofs are applications of the transversality theorem (TT). In our set-up, the parameter space Ξ is not finite dimensional. However, we will always use local perturbations which are polynomial, i.e., finite dimensional. To go from our results to the ones referred to Ξ is a purely technical, and straightforward, matter. We will exploit the TT in several different context. Therefore, we outline here the general procedure. We use (arbitrarily small) perturbations of the vector ν and of the functions (f, b, c). We start with a given function, for instance $f^k(\theta)$, and introduce a polynomial perturbation, setting

$$f^k(\theta; d) = f^k(\theta) + \sum_{v=0}^{V} d_v^k \theta^v,$$

where $d \in D \subset \mathbb{R}^V$, some small open neighborhood of 0. Using TT, we show that a required property holds for all the vectors $d \in D^*$, some open, dense subset of D. This is what we exactly mean saying that a property holds "generically in $f^k(\theta)$" or "modulo a perturbation of $f^k(\theta)$". To use polynomial (hence, finite dimensional) perturbations is convenient, and it does not imply any essential loss of generality with respect to openness and density of the set of functions we restrict ourselves to. We can also perturb in different directions the same function at different vectors in its domain: Pick, for instance, θ_1, θ_2, with $\theta_1 \neq \theta_2$. Choose two open neighborhoods of radius $\varepsilon, V_\varepsilon(\theta_1)$ and $V_\varepsilon(\theta_2)$, such that $clV_{2\varepsilon}(\theta_1) \cap clV_{2\varepsilon}(\theta_2) = \emptyset$. Choose two smooth "bump" functions (with nonnegative values) $\psi_1(\theta)$ and $\psi_2(\theta)$, taking the value 1 on the set $V_\varepsilon(\theta_1)$ $(V_\varepsilon(\theta_2))$ and the value 0 at $\theta \notin clV_{2\varepsilon}(\theta_1)$ $(clV_{2\varepsilon}(\theta_2))$. Functions with these properties exist. Define the perturbed function $f^k(\theta; d) = f^k(\theta) + \psi_1(\theta) d_1 + \psi_2(\theta) d_2$. Evidently, on, say, $V_\varepsilon(\theta_1)$, $\frac{\partial f^k(\theta;d)}{\partial d_1} = 1$ and $\frac{\partial f^k(\theta;d)}{\partial d_1} = 0$ at $\theta \notin clV_{2\varepsilon}(\theta_1)$. In a similar way, and using polynomials, we can arbitrarily (and independently) perturb the derivatives of any order of the functions.

7.3. Existence and regularity of stationary equilibrium

7.3.1. Existence of interior SSE

We start with an outline of the proof. Consider the economy ξ. By A6, T_ξ is the largest value of T such that there is a SSE $\phi(T,\xi)$ with $\Theta^{0e}(\phi(T,\xi);T,\xi)=\Theta^0$. If, at T_ξ, $\Phi(\phi;T,\xi)$ were C^1 and $\det D_\phi\Phi(\phi;T,\xi)\neq 0$, existence of interior SSE, for each $T>T_\xi$ in some open neighborhood of T_ξ, would follow immediately. By IFT, there would be a map $\phi(T,\xi)$ such that $\Phi(\phi(T,\xi);T,\xi)=0$, for T close to T_ξ. Existence of interior SSE would follow, because, by construction, at $T>T_\xi$, $\Theta^{0e}(\phi(T,\xi);T,\xi)\neq\Theta$, and, by continuity, $\Theta^{0e}(\phi(T,\xi);T,\xi)\neq\emptyset$, for T close to T_ξ.

The difficulty is that, at T_ξ, $\Phi(\phi(T,\xi);T,\xi)$ is necessarily non differentiable, because each $\theta^*\in G_\phi^{-1}(0)$ is either on the boundary of $[\theta_\ell,\theta_h]$ or, worst, an interior minimum of $G(\theta,\phi(T,\xi);T,\xi)$, so that $\frac{\partial G}{\partial\theta}=0$. In both cases, $\Phi(\phi(T,\xi);T,\xi)$ fails to be differentiable. If $G_\phi^{-1}(0)$ is always a singleton, at T_ξ, $G_\phi^{-1}(0)$ is a boundary point and it is easy to sidestep this problem. The hard case is when, at T_ξ, $G_\phi^{-1}(0)\in(\theta_\ell,\theta_h)$. For this case, we show that, for T sufficiently close to T_ξ, and an appropriate perturbation of ξ, there is $\widehat{\xi}$ arbitrarily close to ξ, such that there is a SSE $\phi(T,\widehat{\xi})$, $\Phi(\phi(T,\widehat{\xi});T,\widehat{\xi})$ is C^1 and $\det D_\phi\Phi(.)\neq 0$. Existence of this SSE follows by continuity of the maps $G(.)$ and $\Phi(.)$ at $\left(\phi\left(T,\xi_\xi\right),T_\xi\right)$. This is a much weaker condition than differentiability (plus $\frac{\partial G}{\partial\theta}\neq 0$). Still, it is not necessarily satisfied. We overcome this problem using TT.

We start proving that there are economies satisfying A6.

Lemma A1. Under A1-A5, let $v^e=v^{ne}$, $q^e(\phi)=q^{ne}(\phi)$, at each $\phi^e=\phi^{ne}$, and

$$\frac{\int_{\Theta^0}\alpha(\vartheta)\left(f^e(\vartheta)-b^e(\vartheta)\right)d\vartheta}{\int_{\Theta^0}\alpha(\vartheta)\,d\vartheta}>\frac{\int_{\Theta^0}\left(1-\alpha(\vartheta)\right)\left(f^{ne}(\vartheta)-b^{ne}(\vartheta)\right)d\vartheta}{\int_{\Theta^0}\left(1-\alpha(\vartheta)\right)d\vartheta}.$$

Then, for each $\xi\in\Xi$, there is T_ξ such that there is a SSE $\phi(T,\xi)$ with $\Theta^{0e}(\phi(T,\xi);T,\xi)=\Theta^0$ if and only if $T\in(0,T_\xi]$

Proof. Fix $\xi\in\Xi$. Consider the artificial economy with fixed set $\Theta^{0e}=\Theta^0$. The two labor markets are then independent and, evidently, in each one there is a unique SSE ϕ^{k*}. We just need to show that, for T small enough, this is a SSE of the actual economy, i.e., that $G(\theta,\phi^*;T,\xi)\geq 0$, for each θ. By assumption, $F^e(\phi^*;T,\xi)>F^{ne}(\phi^*;T,\xi)$, while $v^e=v^{ne}$ and $q^e(\phi)=q^{ne}(\phi)$, at each $\phi^e=\phi^{ne}$. Therefore, $\phi^{e*}>\phi^{ne*}$, because $\frac{\partial A^k}{\partial\phi^k}>0$, for each k. Consider $G(\theta,\phi^*;T,\xi)$, given by eq. (8). Under A4, if $\phi^e=\phi^{ne}$, the term in square brackets is positive. By direct computation, for each θ, $\frac{\partial G}{\partial\phi^e}>0$. By assumption, $\alpha(\theta)>0$, for each θ. Hence, the first term of $G(\theta,\phi^*;T,\xi)$ is strictly positive for each θ. For $T=0$, the second term is nil. Therefore, for T small enough, $G(\theta,\phi^*;T,\xi)\geq 0$ for each θ, and ϕ^* is a SSE. Evidently, for T large enough, $G(\theta,\phi^*;T,\xi)<0$, for each θ. Hence, the set of values of $T\in\mathbb{R}_{++}$ such that the given pair ϕ^* is a SSE (and $\Theta^e(\phi^*;T,\xi)=\Theta$) is bounded above. Given that $G(\theta,\phi^*;T,\xi)$ is continuous in T, there is

$$T_\xi=\max\left\{T\in(0,\infty)\,|\phi^*\text{ is a SSE at }T\text{ and }\Theta^{0e}(\phi^*;T,\xi)=\Theta\right\},$$

and ϕ^* is a SSE with $\Theta^{0e}(\phi^*;T,\xi)=\Theta$ if and only if $T\leq T_\xi$. ∎

Given an economy $\xi\in\Xi$, let ϕ_ξ be its SSE for $\Theta^{0e}=\Theta^0$. Clearly, ϕ_ξ is unique.

We split the proof of Thm. 1 into several steps. First, we show that, generically, given ξ, at (ϕ_ξ, T_ξ), $G^{-1}_{(\phi_\xi, T_\xi)}(0)$ is a discrete set. Next, we fix (ϕ_ξ, T_ξ) of a given ξ, and show that there is an economy $\widehat{\xi}$ (arbitrarily close to ξ) and some $T_{\widehat{\xi}} > T_\xi$ such that ϕ_ξ is an interior SSE of $\widehat{\xi}$, given $T_{\widehat{\xi}}$. The third step is to show that $\text{rank} D_\phi \Phi(\phi; T_{\widehat{\xi}}, \widehat{\xi}) = 2$, at $\phi = \phi_\xi$. Then, Thm. 1 follows by the IFT.

First, given ξ, we can (locally) restrict the analysis to $\phi \in \digamma_\xi \subset \mathbb{R}^2_{++}$, where $\digamma_\xi$ is a smooth, compact manifold without boundary.

Let $\underline{\phi}^k = \frac{1}{2}\left\{\phi^k | A^k(\phi^k; T, \xi) \ = \min_\theta \left(f^k(\theta) - b^k(\theta)\right)\right\}$, and
$\overline{\phi}^k = 2\left\{\phi^k | A^k(\phi^k; T, \xi) = \max_\theta \left(f^k(\theta) - b^k(\theta)\right)\right\}$.

By A3, $\min_\theta \left(f^k(\theta) - b^k(\theta)\right) > 0$, for each k. By compactness of $[\theta_\ell, \theta_h]$ and continuity of (f, b), there is $\theta \in \arg\max_\theta \left(f^k(\theta) - b^k(\theta)\right)$. For each k, the function $A^k(\phi^k; T, \xi)$ is continuous and strictly increasing. Moreover, given A4, $A^k(\mathbb{R}_{++}; T, \xi) = \mathbb{R}_{++}$. Hence, the pair $(\underline{\phi}^k, \overline{\phi}^k)$ exists and is unique, for each k. Evidently, if a SSE ϕ^* exists, it must be $\phi^* \in int \prod_k [\underline{\phi}^k, \overline{\phi}^k]$. Hence, we can take $\phi \in \digamma_\xi \subset \mathbb{R}^2_{++}$, where $\digamma_\xi$ has the properties stated above and contains $\prod_k [\underline{\phi}^k, \overline{\phi}^k]$.

Fact 1.1. Under A1-A6, there is an open, dense subset $\Xi' \subset \Xi$, such that, for each $\xi \in \Xi'$, $G^{-1}_{(\phi,T)}(0)$ is a discrete set at each (ϕ, T).

Proof of Fact 1.1. Fix $\Theta^{0e} = \Theta^0$. Consider the map $\Psi : [\theta_\ell, \theta_h] \times \mathbb{R}^3_{++} \times \Xi \to \mathbb{R}^5$, defined by

$$\Psi(\theta, \phi; T, \xi) = \begin{pmatrix} \Phi(\theta, \phi; T, \xi) \\ G(\theta, \phi; T, \xi) \\ \frac{\partial G(\theta, \phi; T, \xi)}{\partial \theta} \\ \frac{\partial^2 G(\theta, \phi; T, \xi)}{\partial \theta^2} \end{pmatrix},$$

and replace the function $c(\theta)$ with $c(\theta; d) = c(\theta) + rd_0^c + rd_1^c\theta + rd_2^c\theta^2$, where $d \in D$, a sufficiently small, open subset of $\mathbb{R}^3$, with $0 \in D$. Assume that $\Psi \pitchfork 0$. Then, by TT, except for a null subset of Ξ, $\Psi_\xi \pitchfork 0$. Given that $\Psi_\xi : [\theta_\ell, \theta_h] \times \mathbb{R}^3_{++} \to \mathbb{R}^5$, this implies that $\Psi_\xi^{-1}(0) = \emptyset$, i.e., that, whenever $\Phi_\xi(\theta, \phi; T) = 0$, at each $\theta \in G_\xi^{-1}(0)$, $\frac{\partial^n G_\xi(\theta, \phi; T)}{\partial \theta^n} \neq 0$, for at least one $n \in \{1, 2\}$. Hence, $\theta \in G_\xi^{-1}(0)$ is neither a degenerate local extremum, nor an inflexion point. To show that $\Psi \pitchfork 0$, consider $D_{(v,d)}\Psi(\theta, \phi; T, \xi) =$

$$\begin{bmatrix} v^e & v^{ne} & rd_0^c & rd_1^c & rd_2^c \\ -\frac{r+\beta\pi^e(\phi^e)}{(1-\beta)q^e(\phi^e)} & 0 & 0 & 0 & 0 \\ 0 & -\frac{r+\beta\pi^e(\phi^e)}{(1-\beta)q^e(\phi^e)} & 0 & 0 & 0 \\ 0 & 0 & \left(1-e^{rT}\right) & \left(1-e^{rT}\right)\theta & \left(1-e^{rT}\right)\theta^2 \\ 0 & 0 & 0 & \left(1-e^{rT}\right) & 2\left(1-e^{rT}\right)\theta \\ 0 & 0 & 0 & 0 & 2\left(1-e^{rT}\right) \end{bmatrix},$$

which has full rank (the first row reports the variables we are differentiating with respect to). Hence, by TT, except for a null subset of D, D^1, $\Psi_\xi \pitchfork 0$. Restricting the analysis to $\phi \in \digamma_\xi$, and T lying in some compact, smooth manifold in $\mathbb{R}_+$, we obtain that $D \backslash D^1$ is open and dense. Going from polynomial perturbations to the set Ξ, we establish the result. ■

Pick an economy $\xi \in \Xi'$, T_ξ, and the associated SSE ϕ_ξ. Evidently, for $\varepsilon > 0$ sufficiently small,

a) if $\theta_\ell \in G_c^{-1}(0)$, $\frac{\partial G\left(\theta,\phi_\xi;T_\xi,\xi\right)}{\partial\theta}|_{\theta\in(\theta_\ell,\theta_\ell+\varepsilon)} > 0$;

b) if $\theta_h \in G_{(\phi_\xi,T_\xi)}^{-1}(0)$, $\frac{\partial G\left(\theta,\phi_\xi;T_\xi,\xi\right)}{\partial\theta}|_{\theta\in(\theta_h-\varepsilon,\theta_h)} < 0$;

c) if $\theta' \in G_{(\phi_\xi,T_\xi)}^{-1}(0) \cap (\theta_\ell,\theta_h)$, $\frac{\partial G\left(\theta,\phi_\xi;T_\xi,\xi\right)}{\partial\theta}|_{\theta\in(\theta'-\varepsilon,\theta')} < 0$ and $\frac{\partial G\left(\theta,\phi_\xi;T_\xi,\xi\right)}{\partial\theta}|_{\theta\in(\theta',\theta'+\varepsilon)} > 0$, i.e., θ' is a local minimum.

Fact 1.2. Under A1-A6, there is a dense subset $\Xi'' \subset \Xi'$ such that, for each $\xi \in \Xi''$, there is an interior SSE $\phi(T,\xi)$, for some T. Moreover, at such a SSE, $\frac{\partial G(\theta,\phi(T,\xi);T,\xi)}{\partial\theta}|_{\theta=\theta_m} \neq 0$, at each $\theta_m \in G_{(\phi_\xi,T_\xi)}^{-1}(0)$, $G(\theta_\ell,\phi(T,\xi);T,\xi) \neq 0$, and $G(\theta_h,\phi(T,\xi);T,\xi) \neq 0$.

Proof of Fact 1.2. Take $\xi \in \Xi'$. Fix ϕ_ξ and pick any $T > T_\xi$. $G\left(\theta,\phi_\xi;T,\xi\right)$ is clearly strictly decreasing in T. Therefore, $\Theta^{0e}(\phi_\xi,T;\xi)$ is a proper subset of Θ^0, and, by continuity, nonempty, for each T sufficiently close to T_ξ, $T > T_\xi$. Moreover, for $T > T_\xi$ and sufficiently close to T_ξ, at (ϕ_ξ,T,ξ), $\#G_{(\phi_\xi,T,\xi)}^{-1}(0) \leq 2\#G_{(\phi_\xi,T_\xi,\xi)}^{-1}(0)$, so that $G_{(\phi_\xi,T,\xi)}^{-1}(0)$ is a discrete set, and $\frac{\partial G\left(\theta,\phi_\xi;T,\xi\right)}{\partial\theta} \neq 0$, for each $\theta \in G_{(\phi_\xi,T_\xi,\xi)}^{-1}(0)$. Also, we can pick T such that $G\left(\theta_\ell,\phi_\xi;T,\xi\right) \neq 0$ and $G\left(\theta_h,\phi_\xi;T,\xi\right) \neq 0$. Finally, observe that $(a-c)$ above imply that the correspondence $\Theta^e(\phi_\xi;T,\xi)$ is continuous in T along sequences $\{T^n\}_{n=1}^\infty$, $T^n \geq T_\xi$, for each n (we loose this property when at $\left(\phi_\xi,T_\xi\right)$, $G_{(\phi_\xi,T_\xi,\xi)}^{-1}(0)$ contains an interval). Evidently the properties just established hold for each ξ in some open, dense set $\Xi'' \subset \Xi'$: at each $T > T_\xi$, T close enough to T_ξ, $\Theta^{0e}(\phi_\xi;T,\xi) \neq \Theta^0$ and $\Theta^{0e}(\phi_\xi;T,\xi) \neq \emptyset$. Evidently, the given ϕ_ξ is a SSE at T_ξ, while it is not necessarily a SSE at $T > T_\xi$.

We now perturb the parameter v so that, in the new economy $\widehat{\xi}$, ϕ_ξ is a SSE at some $T > T_\xi$. Given that $G(\theta,\phi;T,\xi)$ does not depend upon v, changes in this parameter have no effect on the set $\Theta^{0e}(\phi_\xi;T,\xi)$. It is easy to check that $F^k\left(\phi_\xi;T,\xi\right)$ is a continuous function of (T,ξ), given that $\mu(\Theta^{k\alpha}(\phi_\xi;T,\xi))$ is locally bounded away from zero. Hence, for $T \to T_\xi$, $\Phi'\left(\phi_\xi;T,\xi\right) \to 0$. Given that $A^k\left(\phi_\xi;T,\xi\right) \neq 0$ (and is $T-$invariant), given any $\varepsilon > 0$, for T sufficiently close to T_ξ, there is $v' \in B_\varepsilon(v)$ such that $\Phi(\phi_\xi;T,(\xi^{\backslash},v')) = 0$. ■

Proof of Theorem 1. Pick $\xi \in \Xi''$, and any T such that, at the associated interior SSE, $\frac{\partial G(\theta,\phi(T,\xi);T,\xi)}{\partial\theta}|_{\theta_m} \neq 0$, at each $\theta_m \in G_{(\phi_\xi,T,\xi)}^{-1}(0) \cap (\theta_\ell,\theta_h)$, $G(\theta_\ell,\phi(T,\xi);T,\xi) \neq 0$ and $G(\theta_h,\phi(T,\xi);T,\xi) \neq 0$. By the IFT, there is a collection of C^1 functions, $\theta_m(\phi;T,\xi)$ locally describing the set $G_{(\phi_\xi,T,\xi)}^{-1}(0)$, i.e., the boundary points of the set $\Theta^e(\phi;T,\xi)$. Therefore, $\Phi(\phi;T,\xi)$ is C^1.

To conclude, all we need is to show that these properties imply that there is also a dense subset $\Xi^\circ \subset \Xi$ such that $D_\phi\Phi(\phi;T,\xi)$ has full rank at an interior SSE. Pick any economy ξ'' in the dense set Ξ'' constructed in Fact 1.2. Let ϕ'' be the associated interior SSE at T''.

The strategy of the proof is the following: First, we show that (locally) arbitrary changes in the pair (v^e,v^{ne}), call them $d_v = (d_v^e,d_v^{ne})$, can be compensated by appropriate changes in the production functions, call them $d_f = (d_f^e(d_v),d_f^{ne}(d_v))$, so that ϕ'' is still a SSE in the new economy at T''. Next, we show that we can pick a d_v arbitrarily small and such that, in the new economy, ξ°, arbitrarily close to ξ'', $\det D_\phi\Phi_{d_v}(\phi;T'',\xi^\circ) \neq 0$ at ϕ''. By definition of density, this implies that there is a dense subset $\Xi^\circ \subset \Xi''$ such that, for each $\xi^\circ \in \Xi^\circ$, there is an interior SSE at some $T(\xi^\circ)$ with $\det D_\phi\Phi(\phi;T(\xi^\circ),\xi^\circ) \neq 0$. By iterated application of IFT, there is an open ball $V(\xi^\circ)$ such that, for each $\xi' \in V(\xi^\circ)$, there is an open ball $V(T(\xi^\circ))$ such

that, at each $T \in V(T(\overset{\circ}{\xi}))$, the economy ξ' has an interior SSE with non-zero determinant. This proves the theorem.

To conclude, we need to construct an economy $\overset{\circ}{\xi}$ with the stated properties, for each $\xi \in \Xi''$. Pick θ^* such that $G(\theta^*, \phi_\xi; T_\xi, \xi) > 0$, and pick some open ball $V_\varepsilon(\theta^*)$ such that, for each $\theta \in clV_{2\varepsilon}(\theta^*)$, $G(\theta, \phi_\xi; T_\xi, \xi) > 0$, and a smooth bump function $\psi(\theta)$. By continuity, there is $V_\varepsilon(\theta^*)$ with the stated properties. For each k, define $f^k(\theta; d) = f(\theta) + \psi(\theta) d_f^k$. By continuity, for ε sufficiently small, this perturbation has no effect on the sets $\Theta^{e\alpha}(\phi_\xi; T_\xi, \xi)$ and $\Theta^{ne\alpha}(\phi_\xi; T_\xi, \xi)$. On the other hand, its effect on the value of $\Phi(\phi_\xi; T_\xi, \xi)$ is

$$\begin{bmatrix} \Delta_f \Phi^e \\ \Delta_f \Phi^{ne} \end{bmatrix} = \begin{bmatrix} \frac{\int_{V_{2\varepsilon}(\theta^*)} \alpha(\theta)\psi(\vartheta)d\vartheta}{\mu\left(\Theta^{e\alpha}(\phi_\xi; T_\xi, \xi)\right)} d_f^e \\ \frac{\int_{V_{2\varepsilon}(\theta^*)} (1-\alpha(\theta))\psi(\vartheta)d\vartheta}{\mu\left(\Theta^{ne\alpha}(\phi_\xi; T_\xi, \xi)\right)} d_f^{ne} \end{bmatrix}.$$

Evidently,

$$\begin{bmatrix} \Delta_v \Phi^e \\ \Delta_v \Phi^{ne} \end{bmatrix} = \begin{bmatrix} -\frac{A^e(\phi^e; T_\xi, \xi)}{v^e} d_v^e \\ -\frac{A^{ne}(\phi^{ne}; T_\xi, \xi)}{v^{ne}} d_v^{ne} \end{bmatrix}.$$

Hence, to preserve the equilibrium, it must be

$$\begin{bmatrix} d_f^e \, (d_v^e) \\ d_f^{ne} \, (d_v^{ne}) \end{bmatrix} = \begin{bmatrix} \frac{A^e(\phi^e; T_\xi, \xi)\mu\left(\Theta^{e\alpha}(\phi_\xi; T_\xi, \xi)\right)}{v^e \int_{V_{2\varepsilon}(\theta^*)} \alpha(\vartheta)\psi(\vartheta)d\vartheta} d_v^e \\ \frac{A^{ne}(\phi^{ne}; T_\xi, \xi)\mu\left(\Theta^{ne\alpha}(\phi_\xi; T_\xi, \xi)\right)}{v^{ne} \int_{V_{2\varepsilon}(\theta^*)} (1-\alpha(\vartheta))\psi(\vartheta)d\vartheta} d_v^{ne} \end{bmatrix}.$$

Let's define

$$\begin{aligned} \Phi\left((\phi; T, \xi), d_f, d_v\right) &\equiv \begin{pmatrix} F^e(.) - A^e(.) + \Delta_f \Phi^e d_f^e + \Delta_v \Phi^e d_v^e \\ F^{ne}(.) - A^{ne}(.) + \Delta_f \Phi^{ne} d_f^{ne} + \Delta_v \Phi^{ne} d_v^{ne} \end{pmatrix} \\ &\equiv \Phi(\phi; T, \xi) + \Delta\left(d_f, d_v\right). \end{aligned}$$

By direct computation, $D_\phi \Delta\left(d_f, d_v\right) \equiv$

$$-\begin{bmatrix} \frac{\Delta_f \Phi^e}{\mu(\Theta^{e\alpha}(.))} \frac{\partial \mu(\Theta^{e\alpha}(.))}{\partial \phi^e} d_f^e + \frac{1}{v^e} \frac{\partial A^e}{\partial \phi^e} d_v^e & \frac{\Delta_f \Phi^e}{\mu(\Theta^{e\alpha}(.))} \frac{\partial \mu(\Theta^{e\alpha}(.))}{\partial \phi^{ne}} d_f^e \\ \frac{\Delta_f \Phi^{ne}}{\mu(\Theta^{ne\alpha}(.))} \frac{\partial \mu(\Theta^{ne\alpha}(.))}{\partial \phi^e} d_f^{ne} & \frac{\Delta_f \Phi^{ne}}{\mu(\Theta^{ne\alpha}(.))} \frac{\partial \mu(\Theta^{ne\alpha}(.))}{\partial \phi^{ne}} d_f^{ne} + \frac{1}{v^{ne}} \frac{\partial A^{ne}}{\partial \phi^{ne}} d_v^{ne} \end{bmatrix}$$

and, substituting into it the vector $d_f\left(d_v\right)$, $D_\phi \Delta\left(d_f(d_v), d_v\right) \equiv$

$$-\begin{bmatrix} \left(\frac{A^e(.)}{\mu(\Theta^{e\alpha}(.))v^e} \frac{\partial \mu(\Theta^{e\alpha}(.))}{\partial \phi^e} + \frac{1}{v^e} \frac{\partial A^e}{\partial \phi^e}\right) d_v^e & \frac{A^e(.)}{v^e \mu(\Theta^{e\alpha}(.))} \frac{\partial \mu(\Theta^{e\alpha}(.))}{\partial \phi^{ne}} d_v^e \\ \frac{A^{ne}(.)}{\mu(\Theta^{ne\alpha}(.))v^{ne}} \frac{\partial \mu(\Theta^{ne\alpha}(.))}{\partial \phi^e} d_v^{ne} & \left(\frac{A^{ne}(.)}{\mu(\Theta^{ne\alpha}(.))v^{ne}} \frac{\partial \mu(\Theta^{ne\alpha}(.))}{\partial \phi^{ne}} + \frac{1}{v^{ne}} \frac{\partial A^{ne}}{\partial \phi^{ne}}\right) d_v^{ne} \end{bmatrix},$$

where $\frac{\partial A^k}{\partial \phi^k} = \frac{1}{1-\beta} \frac{q^k(\phi) \frac{\partial \pi^k}{\partial \phi^k} - (r + \beta\pi^e(\phi^e)) \frac{\partial q^k}{\partial \phi^k}}{q^k(\phi)^2} > 0$, for each k. Given the results above, without any loss of generality, $\frac{\partial G}{\partial \theta}|_{\theta=\theta_m} \neq 0$, at each $\theta_m \in G^{-1}_{(\phi_\xi, T, \xi)}(0)$. Then, using (2), and the function $\chi(\theta_m)$, such that $\chi(\theta_m) = 1$ if θ_m is the lower bound of an interval $[\theta_m, \theta_{m+1}] \subset \Theta^{0e}$, $\chi(\theta_m) = 2$ otherwise,

$$\frac{\partial \mu\left(\Theta^{e\alpha}(.)\right))}{\partial \phi^k} = e^{-\gamma T} \sum_{m=1}^{M} \alpha(\theta_m) (-1)^{\chi(\theta_m)} \frac{\partial \theta_m}{\partial \phi^k},$$

and

$$\frac{\partial \mu\left(\Theta^{ne\alpha}(.)\right)}{\partial \phi^k} = e^{-\gamma T} \sum_{m=1}^{M} \left(1 - e^{\gamma T} - \alpha(\theta_m)\right) (-1)^{\chi(\theta_m)} \frac{\partial \theta_m}{\partial \phi^k}.$$

Therefore,

$$\frac{\partial \mu\left(\Theta^{e\alpha}(.)\right)}{\partial \phi^{e}} \frac{\partial \mu\left(\Theta^{ne\alpha}(.)\right)}{\partial \phi^{ne}} = \frac{\partial \mu\left(\Theta^{e\alpha}(.)\right)}{\partial \phi^{ne}} \frac{\partial \mu\left(\Theta^{ne\alpha}(.)\right)}{\partial \phi^{e}}.$$

Moreover, as shown in Prop. 1, $\frac{\partial \mu\left(\Theta^{k\alpha}(.)\right)}{\partial \phi^{k}} > 0$, and $\frac{\partial \mu\left(\Theta^{k\alpha}(.)\right)}{\partial \phi^{k'}} < 0$. Hence,

$$\begin{aligned} \frac{v^{e}v^{ne}}{d_{v}^{e}d_{v}^{ne}} \det D_{\phi}\Delta\left(.\right) &= \frac{\partial A^{e}}{\partial \phi^{e}} \frac{\partial A^{ne}}{\partial \phi^{ne}} + \frac{\partial A^{e}}{\partial \phi^{e}} \frac{A^{ne}\left(.\right)}{\mu\left(\Theta^{ne\alpha}(.)\right)} \frac{\partial \mu\left(\Theta^{ne\alpha}(.)\right)}{\partial \phi^{ne}} \\ + \frac{\partial A^{ne}}{\partial \phi^{ne}} \frac{\partial \mu\left(\Theta^{e\alpha}(.)\right)}{\partial \phi^{e}} \frac{A^{e}\left(.\right)}{\mu\left(\Theta^{e\alpha}(.)\right)} &> 0. \end{aligned}$$

We conclude showing that, for an appropriate choice of d_v, this implies that $\det D_{\phi}\Phi\left((\phi;\xi),d_{f},d_{v}\right) \neq 0$. Consider any matrix

$$B = \begin{bmatrix} a - a_{1}d_{v}^{e} & b - b_{1}d_{v}^{e} \\ c - c_{1}d_{v}^{ne} & e - e_{1}d_{v}^{ne} \end{bmatrix},$$

with $\det B = (ae - bc) + (bc_1 - ae_1)\, d_v^{ne} + (cb_1 - ea_1)\, d_v^{e} + (a_1e_1 - b_1c_1)\, d_v^{e}d_v^{ne}$. Assume that $(ae - bc) = 0$ (otherwise there is nothing to show) and that $(a_1e_1 - b_1c_1) \neq 0$. If $(bc_1 - ae_1) = (cb_1 - ea_1) = 0$, there is nothing else to show. Otherwise, pick any d_v^e such that $[(bc_1 - ae_1) + (a_1e_1 - b_1c_1)\, d_v^{e}] \neq 0$ (this can be done because $(a_1e_1 - b_1c_1) \neq 0$). Then, $\det B \neq 0$ whenever

$$d_{v}^{ne} \neq \frac{-\left(cb_{1} - ea_{1}\right) d_{v}^{e}}{\left(bc_{1} - ae_{1}\right) + \left(a_{1}e_{1} - b_{1}c_{1}\right) d_{v}^{e}}.$$

Using the notation introduced above, $\det D_{\phi}\Delta\left(d_{f}(d_{v}),d_{v}\right) \neq 0$ means $(a_1e_1 - b_1c_1) \neq 0$. Then, we just pick a pair (d_v^e, d_v^{ne}) satisfying the last two inequalities (so that $\det D_{\phi}\Phi\left((\phi;\xi),d_{f},d_{v}\right) \neq 0$) and sufficiently small, so that the economy ξ° so obtained is sufficiently close to the original economy ξ. ■

7.3.2. Generic regularity of SSE

We start with a preliminary result.

Fact 2.1: Given T, there is an open, dense set of economies $\Xi' \subset \Xi$, such that, at each $\phi \in \digamma_{\xi} \subset \mathbb{R}^{2}_{++}$, $\digamma_{\xi}$ compact, $G_{\phi}^{-1}\left(0\right)$ is either empty, or it contains a finite number of isolated points.

Proof of Fact 2.1: It is similar to the one of Fact 1.1 above. Hence, we just point out how the argument must be modified. Given that we have one less variable (T is now fixed) and two less eqs. (the Proposition holds for all $\phi \in \digamma_{\xi}$), we obtain a somewhat weaker results, because we cannot rule out inflexion points, but only degenerate extrema. Define the system of equations

$$\Psi\left(\theta,\phi;\xi\right) \equiv \begin{pmatrix} G\left(\theta,\phi;\xi\right) \\ \frac{\partial G}{\partial \theta} \\ \frac{\partial^{2} G}{\partial \theta^{2}} \\ \frac{\partial^{3} G}{\partial \theta^{3}} \end{pmatrix} = 0,$$

$\Psi : [\theta_{\ell},\theta_{h}] \times \digamma_{\xi} \times V(\xi^{\circ}) \to \mathbb{R}^{4}$, and $c(\theta;d) = c(\theta) + \sum_{j=0}^{3} r d_{j}\theta^{j}$, for $d = (d_0,\ldots,d_3)$ in some open set $D \subset \mathbb{R}^{4}$, such that $0 \in D$, and $c(\theta;d) > 0$, for each $\theta \in [\theta_{\ell},\theta_{h}]$. It is easy to check that $D_{d}\Psi\left(\theta,\phi;\xi\right)$ is a full rank matrix. Hence, $\Psi \pitchfork 0$, and by the TT, there is a an open, dense set $D' \subset D$, of full Lebesgue measure, such that $\Psi_{\xi}^{-1} \pitchfork 0$, for each $d \in D'$. Given that Ψ_{ξ}^{-1} maps a subset of $\mathbb{R}^3$ into $\mathbb{R}^4$, $\Psi_{\xi}^{-1}\left(0\right) = \emptyset$. Therefore, at each $\theta \in G_{d}^{-1}\left(0\right)$, either $\frac{\partial G_d}{\partial \theta} \neq 0$, or, if $\frac{\partial G_d}{\partial \theta} = 0$, either $\frac{\partial^2 G_d}{\partial \theta^2} \neq 0$ or $\frac{\partial^3 G_d}{\partial \theta^3} \neq 0$ or both. It follows that $G_{d}^{-1}\left(0\right)$ does not contain

any degenerate local extrema and, therefore, that it is a discrete set. Hence, by compactness, $G_d^{-1}(0) \cap [\theta_\ell, \theta_h]$ is a finite set. The same argument used above completes the proof of the claim. ■

Proof of Theorem 2. With an argument similar to the one of Thm. 1, we can show that $\Xi_T \neq \emptyset$ and that it contains an open set Ξ_T°. Now T is fixed and, therefore, omitted from the notation whenever possible.

Consider the intersection $\Xi_T^\circ \cap \Xi'$ (the open and dense set of economies whose properties have been established in Fact 2.1 above). This is an open and dense (in Ξ_T°) set. Given that, for each $\xi \in \Xi'$ and each ϕ in some compact manifold F_ξ, $G_\phi^{-1}(0)$ contains a finite number of isolated points, this holds a fortiori at each SSE of ξ, if any.

To show (i, ii), assume that Ξ_T^{reg} is not dense in $\Xi_T^\circ \cap \Xi'$, i.e., that there is an an open set $V(\xi^\circ) \subset \Xi_T^\circ \cap \Xi' \backslash \Xi_T^{reg}$. We start showing that there is a residual (hence, dense) subset of $V(\xi^\circ) \subset \Xi_T^{reg}$, establishing a contradiction.

Fix $N \in \mathbb{N}$. Pick any collection of N distinct elements of (θ_ℓ, θ_h) with rational coordinates, $\theta(N) = \{\theta_1, ...\theta_n, ..., \theta_N\}$. Define

$$\varepsilon = \min\left\{\min_{n,n'} dist(\theta_n, \theta_{n'}), \min_n dist(\theta_n, \theta_h), \min_n (\theta_n, \theta_\ell)\right\}.$$

Evidently, $\varepsilon > 0$, and $clV_{\frac{2\varepsilon}{5}}(\theta_n) \cap clV_{\frac{2\varepsilon}{5}}(\theta_{n'}) = \emptyset$, for each pair θ_n, $\theta_{n'}$, and θ_ℓ, $\theta_h \notin clV_{\frac{2\varepsilon}{5}}(\theta_n)$, for each n. Consider all the possible partitions of the collection $\theta(N)$ into two (possibly empty) sets, call them $P_s \in \mathcal{P}$. Evidently, the cardinality of $\mathcal{P}$ is finite for each N. Pick a partition $P_s \equiv \left\{P_s^r, P_s^0\right\} \in \mathcal{P}$. Without any loss of generality assume that (modulo a relabelling) $P_s^r = \{\theta_1, ..., \theta_{NR}\}$ and $\theta_1 < ... < \theta_{NR}$. If $\#P_s^r$ is even, define:

1. $\Theta^{0e1}(P_s^r) = [\theta_\ell, \theta_1] \cup [\theta_2, \theta_3] \cup .... \cup [\theta_{NR}, \theta_h]$,

2. $\Theta^{0e2}(P_s^r) = [\theta_1, \theta_2] \cup [\theta_3, \theta_4] \cup .... \cup [\theta_{NR-1}, \theta_{NR}]$.

If odd, define:

3. $\Theta^{0e3}(P_s^r) = [\theta_\ell, \theta_1] \cup [\theta_2, \theta_3] \cup .... \cup [\theta_{NR-1}, \theta_{NR}]$,

4. $\Theta^{0e4}(P_s^r) = [\theta_1, \theta_2] \cup [\theta_3, \theta_4] \cup .... \cup [\theta_{NR}, \theta_h]$.

Use $\zeta, \zeta = 1, ..., 4$, to refer to the indexes above. Redefine the map $\Phi^e(\phi; \xi)$ as $\Phi^e(\theta, \phi; P_s^r, \zeta, \xi) = \left(\frac{\int_{\Theta^{0e\zeta}(P_s^r)} \alpha(\theta)(f^e(\vartheta) - b^e(\vartheta))d\vartheta}{\mu(\Theta^{0e\zeta}(P_s^r))} - A^e(\phi^e; \xi)\right)$, for each ζ. Redefine $\Phi^{ne}(\theta, \phi; P_s^r, \zeta, \xi)$ in a similar way. Set $\theta^a = [\theta_1, ..., \theta_{NR}]$ and $\theta^b = [\theta_{NR+1}, ..., \theta_N]$. Take a fixed $(\widehat{P}_s^r, \widehat{\zeta})$, and define the maps

$$\Phi^E(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) = \begin{bmatrix} \Phi(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) \\ G(\theta, \phi; \xi) \quad \text{at each } \theta_n \in P_s^r \end{bmatrix},$$

$$\Phi^{En'}(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) = \begin{bmatrix} \Phi^E(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) \\ \frac{\partial G(\theta, \phi; \xi)}{\partial \theta}|_{\theta=\theta_{n'}}, \text{ at some } \theta_{n'} \in P_s^r \end{bmatrix}, \text{ for } n' = 1, ..., NR,$$

$$\Psi(\theta, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) = \begin{bmatrix} \Phi^E(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) \\ G(\theta, \phi; \xi), \text{ at each } \theta_n \in P_s^0 \\ \frac{\partial G(\theta, \phi; \xi)}{\partial \theta}|_{\theta=\theta_n}, \text{ at each } \theta_n \in P_s^0 \end{bmatrix},$$

and

$$\Phi^{Ey}(\theta,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi)=\left[\begin{array}{c}\Phi^{E}(\theta^a,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi)\\ G(\theta_y,\phi;\xi)\end{array}\right],\quad \text{for}\quad y=\ell,h.$$

Given that $\Phi^E(\theta^a,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi)$ depends explicitly on both ϕ and the vector θ^a, here it is convenient to define a SSE as a pair $(\theta^{a'},\phi')$. Notice that $(\theta^{a'},\phi')$ is an interior SSE if and only if $\Phi^E(\theta^{a'},\phi';\widehat{P}_s^r,\widehat{\zeta},\xi)=0$ and the associated set $\Theta^{0e\widehat{\zeta}}(\widehat{P}_s^r)$ coincides $a.e.$ with the actual set $\Theta^{0e}(\phi')$. An important difference with respect to the map $\Phi(\theta,\phi;\xi)$ considered above is that $\Phi(\theta,\phi;\xi)$ may fail to be C^1 (because $\frac{\partial\theta_n}{\partial\phi}$ may not exist at some $\theta_n\in G^{-1}(0)$), while $\Phi^E(\theta,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi)$ is always C^1. Also notice that, for each given ζ,

$$\Phi^E:clV_{\frac{2\varepsilon}{5}}(\theta_1)\times...\times clV_{\frac{2\varepsilon}{5}}(\theta_{NR})\times F_{\xi^\circ}\times V(\xi^\circ)\to\mathbb{R}^{2+NR},$$

$$\Phi^{En'}:clV_{\frac{2\varepsilon}{5}}(\theta_1)\times...\times clV_{\frac{2\varepsilon}{5}}(\theta_{NR})\times F_{\xi^\circ}\times V(\xi^\circ)\to\mathbb{R}^{3+NR},\ \text{for each } n'=1,...,NR$$

$$\Psi:clV_{\frac{2\varepsilon}{5}}(\theta_1)\times...\times clV_{\frac{2\varepsilon}{5}}(\theta_{N})\times F_{\xi^\circ}\times V(\xi^\circ)\to\mathbb{R}^{2+NR+2(N-NR)}$$

and

$$\Phi^{Ey}:clV_{\frac{2\varepsilon}{5}}(\theta_1)\times...\times clV_{\frac{2\varepsilon}{5}}(\theta_{NR})\times F_{\xi^\circ}\times V(\xi^\circ)\to\mathbb{R}^{3+NR},\ \text{for each } y,$$

where F_{ξ° can be taken to be a smooth, compact manifold (see proof of Thm. 1) containing all the possible equilibrium values of ϕ (for $\xi\in V(\xi^\circ)$). Given $(\widehat{P}_s^r,\widehat{\zeta})$, assume that all the maps defined above are transversal to 0. Then, for each map, there is an open, dense subset of $V(\xi^\circ)$, call it, for instance, $V_{\Phi^E}(\xi^\circ)$, such that $\Phi_\xi^E\pitchfork 0$ for each $\xi\in V_{\Phi^E}(\xi^\circ)$. For the maps $\Psi_\xi(.)$, $\Phi_\xi^{En'}(.)$, for $n'=1,...,NR$, and $\Phi_\xi^{Ey}(.)$, for each y, this means that there is no solution to the system of eqs., because (given ξ) the dimension of the domain is strictly smaller than the dimension of the range. For the map $\Phi_\xi^E(.)$, $\Phi_\xi^E\pitchfork 0$ means that, for each $\xi\in V_{\Phi^E}(\xi^\circ)$, either $\Phi_\xi^{E-1}(0)=\emptyset$ or $D_{(\theta^a,\phi)}\Phi_\xi^E(.)$ has full rank at each $(\theta^a,\phi)\in\Phi_\xi^{E-1}(0)$. We postpone the proof that the maps defined above are actually transversal to 0. For the time being, just assume so. Define as $V_{(\{\theta_1,...\theta_n,...,\theta_N\},\mathbf{P})}(\xi^\circ)$ the open, dense subset of $V(\xi^\circ)$ obtained as intersection of the (finite) collection of generic subsets of $V(\xi^\circ)$ so obtained.

Repeat the procedure for each ζ and for every $P_s\in\mathcal{P}$. Iterate the procedure for each possible collection $\theta(N)=\{\theta_1,...,\theta_N\}$ with the properties discussed above. Finally, repeat it for each $N\in\mathbb{N}$. We obtain a countable collection of open, dense subsets of $V(\xi^\circ)$. Define as $V'(\xi^\circ)$ the non-empty, residual (hence, dense) set so obtained.

For each $\xi\in V'(\xi^\circ)$, at each SSE (θ,ϕ) such that the vector θ has rational coordinates, properties $(i-ii)$ of the Thm. are satisfied, by construction. Given that any SSE (θ,ϕ) must satisfy $(\theta,\phi)\in G_\xi^{-1}(0)$, for each $\xi\in V'(\xi^\circ)$ (iii) is also satisfied by all SSE such that θ has rational coordinates.

We need to show that the construction can be extended to all the SSE (θ,ϕ). Pick any $\xi'\in V'(\xi^\circ)$ and any interior SSE (θ',ϕ') $\left(\theta'=G^{-1}(0)\cap(\theta_\ell,\theta_h),\phi'\right)$. Given that $\xi'\in\Xi'$, θ' is a finite dimensional vector, with, say, N elements. Partition the set $G^{-1}(0)\cap(\theta_\ell,\theta_h)$ into two vectors $(\theta^{a'},\theta^{b'})$, such that $\frac{\partial G(\phi;\xi)}{\partial\theta}|_{\theta=\theta_n}=0$ if and only if $\theta_n\in\theta^{b'}$. Let

$$\varepsilon'=\min\left\{\min_{n,n'}dist\left(\theta'_n,\theta'_{n'}\right),\min_n dist\left(\theta'_n,\theta_h\right),\min_n\left(\theta_n,\theta'_\ell\right)\right\},$$

where, evidently, $\varepsilon'>0$.

Pick any sequence of N elements with rational coordinates $\{\theta^\nu\equiv(\theta_1^v,...,\theta_N^v)\}_{v=1}^\infty$ such that $\theta^\nu\to\theta'$ (which exists, by definition of $\mathbb{R}^N$). Let $\{\varepsilon^\nu\}_{v=1}^\infty$ be the associated sequence of values of

ε (constructed as above). Evidently, $\varepsilon^{\nu} \to \varepsilon' > 0$. Given that any neighborhood $V_{\frac{\varepsilon'}{5}}(\theta^{'})$ contains a vector with rational coordinates, for v sufficiently large,

$$(\theta^{'}, \phi') \in clV_{\frac{2\varepsilon^{v}}{5}}(\theta_1^{v}) \times ... \times clV_{\frac{2\varepsilon^{v}}{5}}(\theta_N^{v}) \times \digamma_{\xi^{\circ}}.$$

By assumption, and for some $(\widehat{P}_s^r, \widehat{\zeta})$, $\Phi^E(\theta^{a'}, \phi'; \widehat{P}_s^r, \widehat{\zeta}, \xi') = 0$. By construction and TT, given that $\xi' \in V'(\xi^{\circ})$, this implies that $rankD_{(\theta^a,\phi)}\Phi_s^E(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi') = (2 + NR)$ at $(\theta^{a\prime}, \phi')$. Transversality also implies that $\Psi(\theta^{'}, \phi'; \widehat{P}_s^r, \widehat{\zeta}, \xi) = 0$ has no solutions, so that $\theta^{a'} = \theta^{'}$. Finally, given that $\Phi^{Ey}(\theta^{a'}, \phi'; \widehat{P}_s^r, \widehat{\zeta}, \xi) = 0$ has no solution, $G(\theta_y, \phi'; \xi) \neq 0$, for $y = \ell, h$. Hence, the Thm. holds at each interior SE of such a ξ', and, therefore, all the interior SSE of ξ' are regular, i.e., ξ' is a regular economy. In turn, regularity of SSE implies that there is some small open neighborhood of economies $V(\xi')$ such that, for each $\xi \in V(\xi')$, regular SSE are also described by the same collection of smooth functions. Continuity implies that, for $V(\xi')$ sufficiently small, each $\xi \in V(\xi')$ has only regular SSE. Otherwise, we could construct a sequence of non regular equilibria converging to a regular SSE of ξ'. This is impossible. This establishes the Thm. for an open, dense subset of Ξ_T°.

We still have to establish the key fact, i.e., that the maps defined above are actually transversal to 0. We compute derivatives with respect to vacancy costs, v, and to the function $c(\theta)$. Define

$$c(\theta, d^c) = c(\theta) + \sum_{n=1}^{N} \varphi_n(\theta_n)(rd_{0n}^c + rd_{1n}^c\theta) + \varphi_\ell(\theta_\ell)\, rd_{0\ell}^c + \varphi_h(\theta_h)\, rd_{0h}^c$$

where the smooth bump functions $\varphi_n(\theta_n)$ are positive only on the non-intersecting neighborhoods $V_{2\varepsilon}(\theta_n)$. Then, for each given $(\widehat{P}_s^r, \widehat{\zeta})$,

$$D_{(v,rd)}\Phi^E(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) = \begin{bmatrix} dv & rd_{01}^c & \cdots & rd_{0NR}^c \\ D_v\Phi(.) & 0 & 0 & 0 \\ 0 & (1 - e^{rT}) & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1 - e^{rT}) \end{bmatrix},$$

for each $n' = 1, ..., NR$,

$$D_{(v,rd)}\Phi^{En'}(\theta^a, \phi; \widehat{P}_s^r, \widehat{\zeta}, \xi) = \begin{bmatrix} (dv, rd_{01}^c, ..., rd_{0NR}^c) & rd_{1n}^c \\ D_{(v,d)}\Phi^E & \cdots \\ 0 & (1 - e^{rT}) \end{bmatrix},$$

$D_{(v,rd)}\Psi(\theta,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi) =$

$$\begin{bmatrix} (dv, rd_{01}^c, ..., rd_{0NR}^c) & rd_{0NR+1}^c & \cdots & rd_{0N}^c & rd_{1NR+1}^c & \cdots & rd_{1N}^c \\ D_{(v,rd)}\Phi^E & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \left(1-e^{rT}\right) & \vdots & 0 & \left(1-e^{rT}\right)\theta_{NR+1} & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \left(1-e^{rT}\right) & 0 & \vdots & \left(1-e^{rT}\right)\theta_N \\ \vdots & \vdots & \vdots & 0 & \left(1-e^{rT}\right) & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \vdots & \left(1-e^{rT}\right) \end{bmatrix},$$

and

$$D_{(v,rd)}\Phi^{Ey}(\theta,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi) = \begin{bmatrix} (dv, rd_{01}^c, ..., rd_{0NR}^c) & rd_{0y}^c \\ D_{(v,d)}\Phi^E & 0 \\ 0 & \left(1-e^{rT}\right) \end{bmatrix}, \quad \text{for} \quad y = \ell, h.$$

Given that $\text{rank}D_v\Phi(\theta^a,\phi;\widehat{P}_s^r,\widehat{\zeta},\xi) = 2$, all the matrices above have full rank. Hence, all the maps are transversal to 0. ■

7.4. Weak constrained inefficiency

The planner's objective function is given by the stationary value of the expected surplus. Hence, $P'(o,u,\Theta^{0e};\xi) \equiv$

$$\begin{aligned} &\int_0^\infty e^{-\gamma t}\left(\int_{\Theta^{0ne}}\left(\left(1-u^{ne}\right)f^{ne}(\vartheta)+u^{ne}b^{ne}(\vartheta)\right)d\vartheta\right)dt \\ &+\int_T^\infty e^{-\gamma t}\left(\int_{\Theta^{0e}}\left(1-\alpha(\theta)\right)\left(\left(1-u^{ne}\right)f^{ne}(\vartheta)+u^{ne}b^{ne}(\vartheta)\right)d\vartheta\right)dt \\ &+\int_T^\infty e^{-\gamma t}\left(\int_{\Theta^{0e}}\alpha(\theta)\left(\left(1-u^{e}\right)f^{e}(\vartheta)+u^{e}b^{e}(\vartheta)\right)d\vartheta\right)dt \\ &-\int_0^\infty e^{-\gamma t}\left(v^{ne}o^{ne}\int_{\Theta^{0ne}}d\vartheta\right)dt-\int_T^\infty e^{-\gamma t}\left(v^{ne}o^{ne}\int_{\Theta^{0e}}\left(1-\alpha(\theta)\right)d\vartheta\right)dt \\ &-\int_T^\infty e^{-\gamma t}\left(v^{e}o^{e}\int_{\Theta^{0e}}\left(1-\alpha(\theta)\right)d\vartheta\right)dt-\int_0^T e^{-\gamma t}\left(\int_{\Theta^{0e}}c(\vartheta)\,d\vartheta\right)dt. \end{aligned} \tag{12}$$

The first three terms are the discounted surplus of the three different groups of people (uneducated, people who invested in education and failed, educated ones). The next three terms are the associated vacancy costs. The last term gives the cumulated direct costs of education.

Using the steady state conditions (7) and (8), $o^k = \phi^k u^k$, and integrating over time, we can rewrite $P'(o,u,\Theta^{0e};\xi)$ as

$$\begin{aligned} P(\phi,\Theta^{0e};\xi) \quad = \quad & \left(e^{-\gamma T}\int_{\Theta^{0e}}T(1,\vartheta,\phi;\xi)\,d\vartheta+\frac{\int_\Theta\left[\pi^{ne}(\phi)f^{ne}(\vartheta)+\gamma b^{ne}(\vartheta)\right]d\vartheta}{\gamma+\pi^{ne}(\phi)}\right) \\ & -e^{-\gamma T}\frac{\gamma v^e\phi^e\int_{\Theta^{0e}}\alpha(\vartheta)\,d\vartheta}{\gamma+\pi^e(\phi)} \end{aligned}$$

$$-\frac{\gamma v^{ne}\phi^{ne}\left[\int_{\Theta\backslash\Theta^{0e}} d\vartheta + e^{-\gamma T}\int_{\Theta^{0e}}\left(1-\alpha\left(\vartheta\right)\right)d\vartheta\right]}{\gamma+\pi^{ne}\left(\phi\right)},$$

where $T\left(1,\vartheta,\phi;\xi\right)$ is the value at $\rho=1$ of the function

$$\begin{aligned} T\left(\rho,\theta,\phi;\xi\right) &= \alpha\left(\theta\right)\frac{\rho\pi^{e}\left(\phi\right)f^{e}\left(\theta\right)+\gamma b^{e}\left(\theta\right)}{\gamma+\rho\pi^{e}\left(\phi\right)}+\left(1-e^{\gamma T}\right)c\left(\theta\right) \\ &\quad+\left(1-e^{\gamma T}-\alpha\left(\theta\right)\right)\frac{\rho\pi^{ne}(\phi^{ne})f^{ne}\left(\theta\right)+\gamma b^{ne}\left(\theta\right)}{\gamma+\rho\pi^{ne}\left(\phi\right)}. \end{aligned}$$

Proof of Theorem 3. To avoid possible misunderstandings, we will denote $\Omega^{0e},\Omega^{0e\alpha}$ etc. the sets chosen by the planner. Define the maps $F^{k}\left(\Omega^{k\alpha};\xi\right)$ in the obvious way.

Consider the set $\Xi^{''}$ of the economies such that every regular interior SSE allocation, if it exists, is not WCO. We start showing that $\Xi^{''}$ is dense in Ξ. Assume, by contradiction, that there is some open set $V\left(\xi^{\circ}\right)\subset\Xi\backslash\Xi^{''}$. Without loss of generality (in view of Thm. 2), we can assume that each $\xi\in V\left(\xi^{\circ}\right)$ is a regular economy with SSE described by a collection of smooth functions $\left(\phi^{1}\left(\xi\right),...,\phi^{N}\left(\xi\right)\right)$. Consider the SSE described by $\phi^{1}\left(\xi\right)$. Regularity also implies that, for each $\xi\in V\left(\xi^{\circ}\right)$, the correspondence $G_{\xi}^{-1}\left(0\right)$ evaluated at $\left(\phi^{1}\left(\xi\right),\xi\right)$ is described by a finite collection of smooth functions $\left(\theta_{1}\left(\phi^{1}\left(\xi\right),\xi\right),...,\theta_{M}\left(\phi^{1}\left(\xi\right),\xi\right)\right)$. Moreover, at each SSE $G(\theta_{\ell},\phi^{1}\left(\xi\right),\xi)\neq0$ and $G(\theta_{h},\phi^{1}\left(\xi\right),\xi)\neq0$.

To avoid unnecessary problems, it is convenient to restrict further the optimization problem of the planner, by requiring that

1. the set Ω^{0e} has the same structure of the set $\Theta^{e\alpha}$ associated with the SSE,
2. the interior boundary points $\{\theta_{1},...,\theta_{M}\}$ lie in some small non-intersecting open neighborhoods of the SSE boundary points.

For instance, if $\Theta^{e\alpha}=[\theta_{\ell},\theta_{1}^{*}]\cup[\theta_{M}^{*},\theta_{h}]\cup_{2}^{M-1}\left[\theta_{m}^{*},\theta_{m+1}^{*}\right]$, the (modified) planner's optimization problem is

$$\begin{aligned} \left(\phi,\theta_{1},...,\theta_{M}\right) &\in \arg\max\ P(\phi,\theta_{1},...,\theta_{M})\quad \textit{subject to}\quad \Phi_{\Omega^{0e}}\left(\phi;\xi\right)=0 \\ \theta_{m}^{*}-\varepsilon &< \theta_{m}<\theta_{m}^{*}+\varepsilon,\ \text{for each } m. \end{aligned} \tag{13}$$

Evidently, the planner's problem (13) may have no solution (because the constraint is not compact). However, if the SSE is WCO, the SSE vector $(\phi^{*},\theta_{1}^{*},...,\theta_{M}^{*})$ is also a solution to the stated optimization problem. Our approach is to show that the necessary FOCs of (14) are typically violated at a SSE. This immediately implies that the SSE is not a solution to (13) and, a fortiori, that it is not WCO.

The FOCs for an interior solution to the Lagrange problem (13), $max\Lambda\left(\phi,\theta_{1},...,\theta_{m},\delta\right)$, are given by

i. $\frac{\partial\Lambda}{\partial\phi^{k}}=\frac{\partial P}{\partial\phi^{k}}-\delta^{k}\frac{\partial\Phi_{\Omega^{0e}}^{k}(\phi;\xi)}{\partial\phi^{k}}=0$, for each k,

ii. $\frac{\partial\Lambda}{\partial\theta_{m}}=\frac{\partial P}{\partial\theta_{m}}-\sum_{k}\delta^{k}\frac{\partial\Phi_{\Omega^{0e}}^{k}(\phi;\xi)}{\partial\theta_{m}}=0$, for each m,

iii. $\frac{\partial\Lambda}{\partial\delta^{k}}=-\Phi_{\Omega^{0k}}\left(\phi;\xi\right)=0$, for each k,

where $(\delta^e, \delta^{ne}) \in \mathbb{R}^2_+$ is the vector of Lagrange multipliers.

The complete system of equations defining a SSE is given by

$a.$ $\quad \Phi(\phi;\xi) = 0,$

$b.$ $\quad G(\theta_m, \phi;\xi) = 0$, each m.

We now show that, for a generic set of parameters ξ, if $(\phi, (\theta_1, ..., \theta_M))$ solves $(a-b)$, there is no strictly positive vector of Lagrange multipliers such that it also solves $(i-iii)$.

Define

$$\Psi(\phi, (\theta_1, ..., \theta_M), \delta;\xi) \equiv \begin{bmatrix} \frac{\partial \Lambda}{\partial \phi} \\ \Phi_{\Omega^{0e}}(\phi;\xi) \\ \frac{\partial \Lambda}{\partial \theta}|_{\theta=\theta_m}, \quad \text{each } m \\ G(\theta_m, \phi;\xi), \quad \text{each } m \end{bmatrix},$$

where $\Psi : \mathbb{R}^{M+4}_{++} \times V(\xi^\circ) \to \mathbb{R}^{2M+4}$,

Evidently, a WCO allocation must satisfy $\Psi(\phi, (\theta_1, ..., \theta_M), \delta; \xi^\circ) = 0$. Indeed, the first three blocks of $(M+4)$ equations are the FOCs a WCO allocation must satisfy. The last M equations must be satisfied for $(\theta_1, ..., \theta_M)$ to be the set of (local) threshold values at the SSE. Assume that $\Psi \pitchfork 0$. Then, for each ξ in some dense subset of $V(\xi^\circ)$, $\Psi_\xi \pitchfork 0$, which implies that $\Psi_\xi^{-1}(0) = \emptyset$. Hence, our proof reduces to show that $\Psi \pitchfork 0$. Let $G^M(.) = [G(\theta_1, \phi;\xi), ..., G(\theta_M, \phi;\xi)]$. By direct computation, $D_{(\delta,\xi)}\Psi(\phi, (\theta_1, ..., \theta_M), \delta;\xi)$ contains the following submatrix

$$\begin{bmatrix} d\delta^e & d\delta^{ne} & dv^e & dv^{ne} & rdc & d\varphi \\ -\frac{\partial \Phi^e_{\Omega^{0e}}}{\partial \phi^e} & 0 & * & 0 & 0 & 0 \\ 0 & -\frac{\partial \Phi^{ne}_{\Omega^{0e}}}{\partial \phi^{ne}} & 0 & * & 0 & 0 \\ 0 & 0 & \frac{\partial \Phi^e_{\Omega^{0e}}}{\partial v^e} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial \Phi^{ne}_{\Omega^{0e}}}{\partial v^{ne}} & 0 & 0 \\ * & * & * & * & D^2_{(\theta,dc)}\Lambda & D^2_{(\theta,d\varphi)}\Lambda \\ 0 & 0 & 0 & 0 & D_{dc}G^M & D_{d\varphi}G^M \end{bmatrix},$$

where "*" denotes (possibly) non-zero coefficients. The meaning of $(d\delta, dv)$ is clear. The associated columns are linearly independent because $\frac{\partial \Phi^k_{\Omega^{0e}}}{\partial \phi^k} < 0$ and $\frac{\partial \Phi^k_{\Omega^{0e}}}{\partial v^k} < 0$, for each k. The last two variables denote derivatives with respect to parameters affecting the functions $(c(\theta), f^k(\theta), b^k(\theta))$. Pick a collection of M open balls of radius ε centered on $\theta_m, V_\varepsilon(\theta_m)$, such that $clV_{2\varepsilon}(\theta_m) \cap clV_{2\varepsilon}(\theta_{m'}) = \emptyset$, for each pair θ_m and $\theta_{m'}$. Also, pick a collection of smooth bump functions $\psi_m(\theta)$, such that $\psi_m(\theta) = 1$ for $\theta \in V_\varepsilon(\theta_m)$, $\psi_m(\theta) = 0$ for $\theta \notin clV_{2\varepsilon}(\theta_m)$. For the column dc, define $c(\theta; d_c) = c(\theta) + \sum_m \psi_m(\theta) rd_{c_m}$. Given that

$$\begin{aligned} \frac{\partial \Lambda}{\partial \theta_m} &= -\sum_k \delta^k \frac{\partial \Phi^k_{\Omega^{0e}}(\phi;\xi)}{\partial \theta_m} + (-1)^{\chi(\theta_m)} e^{-\gamma T} \\ &\times \left[T(1, \theta_m, \phi;\xi) - \left(\frac{\alpha(\theta_m)\gamma v^e \phi^e}{\gamma + \pi^e(\phi)} + \left(1 - e^{\gamma T} - \alpha(\theta_m)\right) \frac{\gamma v^{ne} \phi^{ne}}{\gamma + \pi^{ne}(\phi)} \right) \right], \end{aligned}$$

we have

$$D^2_{(\theta,dc)}\Lambda = e^{-\gamma T}\left(1 - e^{\gamma T}\right) \begin{bmatrix} (-1)^{\chi(\theta_1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (-1)^{\chi(\theta_M)} \end{bmatrix}.$$

Clearly, $D_{dc}G^M = \left(1-e^{\gamma T}\right)[I]$, where $[I]$ is the $M \times M$ identity matrix. The structure of column dc described above follows immediately, because $\int_{\Theta^{0e}} c(\theta)\, d\theta$ does not directly affect the first four rows of $\Psi_{\theta_1}(\phi, \delta; \xi)$.

Consider now the last column. Pick η sufficiently small, such that $(\theta_m + 8\eta) < \left(\theta_m + \frac{\theta_{m+1}-\theta_m}{2}\right)$, for each m such that θ_m is a lower bound of an interval $[\theta_m, \theta_{m+1}] \subset \Theta^{0e}$, and $(\theta_m - 8\eta) < \left(\theta_m - \frac{\theta_m - \theta_{m-1}}{2}\right)$, for each upper bound θ_m. When θ_m is a lower bound, define $\widehat{\theta}_m \equiv (\theta_m + 5\eta)$. If it is an upper bound, $\widehat{\theta}_m \equiv (\theta_m - 5\eta)$. Bear in mind that, by construction, for each m, $G(\theta, \phi; \xi) > 0$ for each $\theta \in clV_{2\eta}(\widehat{\theta}_m)$, and that the collection $\left\{clV_{2\eta}(\theta_1), ..., clV_{2\eta}(\theta_M), clV_{2\eta}(\widehat{\theta}_1), ..., clV_{2\eta}(\widehat{\theta}_M)\right\}$ is composed by mutually disjoint sets. Given $\{\theta_1, ..., \theta_M\}$ and the associated collection $\left\{\widehat{\theta}_1, ..., \widehat{\theta}_M\right\}$ define, as above, the bump function $\psi_m(\theta)$ $(\widehat{\psi}_m(\theta))$, such that $\psi_m(\theta) = 1$ for $\theta \in V_\eta(\theta_m)$, $\psi_m(\theta) = 0$ for $\theta \notin clV_{2\eta}(\theta_m)$ (that $\widehat{\psi}_m(\theta) = 1$ for $\theta \in V_\eta(\widehat{\theta}_m)$, $\widehat{\psi}_m(\theta) = 0$ for $\theta \notin clV_{2\eta}(\widehat{\theta}_m)$). Define the two perturbations

$$f^e(\theta; d_\varphi) = f^e(\theta) + \sum_m \psi_m(\theta)\, d_{\varphi_m} - \sum_m \widehat{\psi}_m(\theta)\, \zeta_m$$

and

$$b^e(\theta; d_\varphi) = b^e(\theta) - \sum_m \psi_m(\theta)\, d_{\varphi_m}.$$

Fix $\zeta_m = 2\frac{\sum_m \int_{clV_{2\eta}(\theta_m)\cap[\theta_m, \theta_m+1]} \psi_m(\vartheta) d_{\varphi_m} d\vartheta}{\sum_m \int_{clV_{2\eta}(\widehat{\theta}_m)} \widehat{\psi}_m(\vartheta) d\vartheta}$ if θ_m is the lower bound of an interval in Ω^{0e}, $\zeta_m = 2\frac{\sum_m \int_{clV_{2\eta}(\theta_m)\cap[\theta_m-1, \theta_m]} \psi_m(\vartheta) d_{\varphi_m} d\vartheta}{\sum_m \int_{clV_{2\eta}(\widehat{\theta}_m)} \widehat{\psi}_m(\vartheta) d\vartheta}$ if it is an upper bound. By construction, integrating over any interval $[\theta_m, \theta_{m+1}]$ contained in Ω^{0e} the effects cancel out. Therefore, these perturbations have no direct effect on the functions $F^k(\Omega^{k\alpha}; \xi)$, each k, so that $\frac{\partial F^k\left(\Omega^{k\alpha}; \xi\right)}{\partial d_\varphi} = 0$. Given that

$$\begin{aligned}\frac{\partial \Lambda}{\partial \phi^e} &= \frac{\gamma e^{-\gamma T}}{\left(\gamma + \pi^e(\phi)\right)^2} \int_{\Omega^{0e}} \alpha(\vartheta) \left(\left(f^e(\vartheta) - b^e(\vartheta)\right) \frac{\partial \pi^e}{\partial \phi^e} + \left(\phi^e \frac{\partial \pi^e}{\partial \phi^e} - \gamma - \pi^e(\phi) \right) v^e \right) d\vartheta \\ &\quad -\delta^e \frac{\partial \Phi^e_{\Omega^{0e}}(\phi; \xi)}{\partial \phi^e},\end{aligned}$$

$D^2_{(\phi^e, d_\varphi)}\Lambda = 0$, by construction. Evidently, $D^2_{(\phi^{ne}, d_\varphi)}\Lambda = 0$, because $\frac{\partial \Lambda}{\partial \phi^{ne}}$ does not depend upon (f^e, b^e, c). Moreover, given that $\frac{\partial F^k\left(\Omega^{k\alpha}; \xi\right)}{\partial d_\varphi} = \frac{\partial A^k\left(\phi^k; \xi\right)}{\partial d_\varphi} = 0$, $\frac{\partial \Phi^e_{\Omega^{0e}}}{\partial d_\varphi} = \frac{\partial \Phi^{ne}_{\Omega^{0e}}}{\partial d_\varphi} = 0$. Hence, the first four coefficients of the last column are zero. By direct computation,

$$D^2_{(\theta, d_\varphi)}\Lambda = e^{-\gamma T} \frac{\pi^e(\phi) - \gamma}{\gamma + \pi^e(\phi)} \begin{bmatrix} (-1)^{\chi(\theta_1)}\alpha(\theta_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (-1)^{\chi(\theta_M)}\alpha(\theta_M) \end{bmatrix},$$

and

$$D_{d_\varphi} G^M|_{\theta=\theta_m} = \frac{\beta\pi^e(\phi) - \gamma}{\gamma + \beta\pi^e(\phi)} \begin{bmatrix} \alpha(\theta_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha(\theta_M) \end{bmatrix}.$$

Given the structure of $D_{(\delta,\xi)}\Psi(\phi, (\theta_1, ..., \theta_M), \delta; \xi)$, to prove that $\Psi \pitchfork 0$, it suffices to show that the bottom right $2M \times 2M$ matrix has full rank. Divide the first M columns by $\left(1-e^{\gamma T}\right)$, column $(M+1)$ by $\alpha(\theta_1)$, column $(M+2)$ by $\alpha(\theta_2)$, and so on. Finally, divide its first row by $e^{-\gamma T}(-1)^{\chi(\theta_1)}$, the second row by $e^{-\gamma T}(-1)^{\chi(\theta_2)}$ and so on (up to row M). The matrix is now reduced to

$$\begin{bmatrix} I & \frac{\pi^e(\phi)-\gamma}{\gamma+\pi^e(\phi)}[I] \\ I & \frac{\beta\pi^e(\phi)-\gamma}{\gamma+\beta\pi^e(\phi)}[I] \end{bmatrix}.$$

Given that $\beta < 1$, this matrix has obviously full rank $2M$. Hence, $\Psi \pitchfork 0$ and, therefore, by TT, there is a dense subset $D^1 \subset D$ such that $\Psi_\xi^{-1}(0) = \emptyset$. Given that $\Psi_\xi(\theta, \phi, \delta) : \mathbb{R}_{++}^{M+4} \to \mathbb{R}^{2M+4}$, for each economy in this dense set, $\Psi_\xi \pitchfork 0$ implies that $\Psi_\xi(\theta, \phi, \delta) = 0$ has no solution. Regularity of the SSE immediately implies that D^1 is also open. Given that the number of equilibria of $\xi \in V(\xi^\circ)$ is finite, we can iterate the same procedure, obtaining a finite collection of open, dense subsets of D, call them $\{D^1, ..., D^N\}$. Their intersection is also an open, dense subset of D. ■

7.5. Two polar cases

Proof of Proposition 2. To establish complementarity, observe that, by direct computation,

$$\begin{aligned}\frac{\partial G}{\partial \theta} &= \frac{\partial \alpha}{\partial \theta}\left[\frac{\pi^e(\phi)\beta f^e(\theta)}{r+\beta\pi^e(\phi)} - \frac{\pi^{ne}(\phi)\beta f^{ne}(\theta)}{r+\beta\pi^{ne}(\phi)}\right] \\ &\quad + \alpha(\theta)\frac{\pi^e(\phi)\beta}{r+\beta\pi^e(\phi)}\frac{\partial f^e(\theta)}{\partial\theta} + \left(1-e^{rT}-\alpha(\theta)\right)\frac{\pi^{ne}(\phi)\beta}{r+\beta\pi^{ne}(\phi)}\frac{\partial f^{ne}(\theta)}{\partial\theta}.\end{aligned}$$

Multiplying by θ and rearranging, we get

$$\begin{aligned}\theta\frac{\partial G}{\partial\theta} &= \eta_\alpha(\theta)\alpha(\theta)\left[\frac{\pi^e(\phi)\beta f^e(\theta)}{r+\beta\pi^e(\phi)} - \frac{\pi^{ne}(\phi)\beta f^{ne}(\theta)}{r+\beta\pi^{ne}(\phi)}\right] + \\ &\quad \eta_{f^e}(\theta)\left[\alpha(\theta)\frac{\pi^e(\phi)\beta f^e(\theta)}{r+\beta\pi^e(\phi)} + (1-e^{rT}-\alpha(\theta))\frac{\pi^{ne}(\phi)\beta f^{ne}(\theta)}{r+\beta\pi^{ne}(\phi)}\frac{\eta_{f^{ne}}(\theta)}{\eta_{f^e}(\theta)}\right].\end{aligned}$$

Consider any $\theta \in G_\phi^{-1}(0)$. If $\eta_\alpha(\theta) > 0$ the first term is positive. By A3, $\eta_{f^e}(\theta) > 0$ and, when $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)} = 1$, the second term in square brackets is $G(\theta, \phi)$ and it must be nil. Hence, this term is nonnegative as long as $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)} \geq 1$ for each θ. This establishes the first claim. b can be shown similarly. ■

Proof of Proposition 3. Given an interior SSE $\overline{\phi}$, consider the direct effect $\frac{\partial P}{\partial\theta}|_{\theta=\overline{\theta}} \equiv \frac{\partial P}{\partial\overline{\theta}}$, computed at the unique threshold $\overline{\theta} \in G_\phi^{-1}(0)$. Using A7, the direct effect of a change in the threshold $\overline{\theta}$ on total surplus is

$$\begin{aligned}e^{\gamma T}(-1)^{\chi(\overline{\theta})}\frac{\partial P}{\partial\overline{\theta}} &= \left(\frac{\alpha(\overline{\theta})\pi^e(\overline{\phi})f^e(\overline{\theta})}{\gamma+\pi^e(\overline{\phi})} + \frac{\left(1-e^{\gamma T}-\alpha(\overline{\theta})\right)\pi^{ne}(\overline{\phi})f^{ne}(\overline{\theta})}{\gamma+\pi^{ne}(\overline{\phi})}\right) \\ &\quad - \left(\frac{\alpha(\overline{\theta})\gamma v\overline{\phi}^e}{\gamma+\pi^e(\overline{\phi})} + \left(1-e^{\gamma T}-\alpha(\overline{\theta})\right)\frac{\gamma v\overline{\phi}^{ne}}{\gamma+\pi^{ne}(\overline{\phi})}\right).\end{aligned}$$

With complementarity, it is always $\pi^e(\overline{\phi}) > \pi^{ne}(\overline{\phi})$. Hence, the first term in brackets ($T\left(1, \overline{\theta}, \overline{\phi}; \xi\right)$, using the notation introduced above) is always negative, because of two facts. First, $T\left(\rho, \overline{\theta}, \phi; \xi\right) = 0$ if and only if $\frac{(\gamma+\rho\pi^e(\overline{\phi}))(\gamma+\rho\pi^{ne}(\overline{\phi}))}{\rho}T\left(\rho, \overline{\theta}, \overline{\phi}; \xi\right) = 0$. This last equation is linear in ρ, and hence it has a unique solution, $\beta = \rho$. At $\rho = \beta$, if $(\pi^e(\phi) - \pi^{ne}(\phi)) > 0$, $\frac{\partial T\left(\rho, \overline{\theta}, \overline{\phi}; \xi\right)}{\partial\rho} < 0$, so that $T\left(1, \overline{\theta}, \overline{\phi}; \xi\right) < 0$.[23]

Consider now the second term in brackets. Under the maintained assumptions, $\frac{\overline{\phi}^e}{\gamma+\pi^e(\overline{\phi})}$ is an increasing function of $\overline{\phi}^e$, unbounded above. Hence, for $\frac{\pi^e(\overline{\phi})}{\pi^{ne}(\overline{\phi})}$ (i.e., $\frac{\overline{\phi}^e}{\overline{\phi}^{ne}}$) sufficiently large, this term is positive so that $(-1)^{\chi(\overline{\theta})}\frac{\partial P}{\partial\overline{\theta}} < 0$. A sufficient condition to obtain an arbitrarily large ratio $\frac{\overline{\phi}^e}{\overline{\phi}^{ne}}$ at the SSE is to have $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ sufficiently large, for each θ.

[23] Comparing with footnote 18, here $\pi^e(\overline{\phi}) > \pi^{ne}(\overline{\phi})$ always implies $T(1, \overline{\theta}, \overline{\phi}; \xi) < 0$ because $c(\theta) = 0$.

Consider now the indirect effect. Under the maintained assumptions, for $\alpha(\theta)$ sufficiently large for each θ, $(f^{ne}(\overline{\theta}) - F^{ne}(\overline{\theta})) > 0$ for each $\overline{\theta}$, and $\frac{\partial \phi^k}{\partial \overline{\theta}} > 0$, for each k. Therefore, if $(\beta + \eta_{q^k}) \geq 0$, each k,

$$(-1)^{\chi(\overline{\theta})} \frac{\partial P}{\partial \phi^k} \frac{\partial \phi^k}{\partial \overline{\theta}} = (-1)^{\chi(\overline{\theta})} \sum_k \frac{\gamma v \mu(\Theta^{k\alpha}(\overline{\phi}; \xi))(\beta + \eta_{q^k})}{(1-\beta)(\gamma + \pi^k(\phi))} \frac{\partial \phi^k}{\partial \overline{\theta}} < 0,$$

because $\chi(\overline{\theta}) = 1$. Hence, weak constrained inefficiency of SSE is due to overeducation.

If $(\beta + \eta_{q^k}) < 0$, $(-1)^{\chi(\overline{\theta})} \frac{\partial P}{\partial \phi^k} \frac{\partial \phi^k}{\partial \overline{\theta}} > 0$, but its value can be made arbitrarily small (for each $\overline{\theta}$), if $(f^k(\overline{\theta}) - F^k(\overline{\theta}))$ is sufficiently small, for each k and $\overline{\theta}$. A sufficient condition to obtain this property is a sufficiently small value of $\eta_{f^k}(\theta)$, for each k. Hence, for $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ sufficiently large and $\eta_{f^k}(\theta)$ sufficiently small, for each k, b holds. ■

Proof of Proposition 4. Assume that $\beta \geq |\eta_{q^k}|$. Then, the welfare impact of the composition effect is positive. For $\frac{\overline{\phi}^e}{\overline{\phi}^{ne}}$ not "too large" the direct effect is also positive (which requires $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ not "too large" for each θ), so that undereducation holds.

Assume that $\beta < |\eta_{q^k}|$. Then, the welfare impact of the composition effect is negative. Sufficiently small values of $\eta_{f^e}(\theta)$ and $\eta_{f^{ne}}(\theta)$ guarantee that the absolute value of the indirect effect is small. We need a sufficiently large (in absolute value) positive direct effect, for undereducation to hold. This obtains if $\frac{\eta_{f^e}(\theta)}{\eta_{f^{ne}}(\theta)}$ is not too large. ■

References

[1] Acemoglu, D.T., 1996, A microfoundation for social increasing returns in human capital accumulation, Q.J.E, 111, 779-804.

[2] Allen, B., 1985, The existence of rational expectations equilibria in a large economy with noisy price observations, J. Math. Econ., 14, 67-103.

[3] Al-Najjar, N.I., 2004, Aggregation and the law of large numbers in large economies, Games Econ. Behavior, 47, 1-35.

[4] Alòs-Ferrer, C., 1999, Dynamical systems with a continuum of randomly matched agents, J. Econ. Theory, 86, 245-267.

[5] Alòs-Ferrer, C., 2002, Individual randomness in economic models with a continuum of agents, mimeo.

[6] Becker, G.S., 1964, Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, Columbia University Press, New York.

[7] Ben-Porath, Y., 1967, The production of human capital and the life-cycle of earnings," J. Pol. Econ., 75, 352-365.

[8] Booth, A.L., and M. Coles, 2005, Increasing returns to education and the skills under-investment trap, IZA DP No. 1657.

[9] Booth, A.L., and M. Coles, 2007, A microfoundation for increasing returns in human capital accumulation and the under-participation trap, Europ. Econ. Review, 51, 1661-1681

[10] Booth, A.L., M. Coles, and X. Gong, 2006, Increasing returns to education: theory and evidence, The Australian National University, Centre for Economic Policy Research, D.P. 522.

[11] Burdett, K., C. Carrillo-Tudela, and M.G. Coles, 2009, Human Capital Accumulation and Labour Market Equilibrium, IZA DP No. 4215.

[12] Burdett, K. and E. Smith, 1996, Education and matching externalities, in Booth, A.L., and D.J. Snower (eds.), Acquiring Skills: Market Failures, Their Symptoms and Policy Responses, Cambridge University Press, Cambridge, 65-80.

[13] Burdett, K. and E. Smith, 2002, The low skill trap, Europ. Econ. Review, 46, 1439-1451.

[14] Carneiro, P., J.J. Heckman, and E. Vytlacil, 2001, Estimating the return to education when it varies among individuals, mimeo.

[15] Cahuc, P., F. Postel-Vinay, and J.-M. Robin, 2006, Wage bargaining with on-the-job search: Theory and evidence, Econometrica, 74, 323-364.

[16] Charlot, O. and B. Decreuse, 2005, Self-selection in education with matching frictions, Labour Econ., 12, 251-267.

[17] Charlot, O. and B. Decreuse, 2006, Over-education for the rich vs. under-education for the poor: A search-theoretic microfoundation, GRECAM, Document de Travail 2006-2.

[18] Charlot, O., B. Decreuse, and P. Granier, 2005, Adaptability, productivity, and educational incentives in a matching model, Europ. Econ. Review, 49, 1007-1032.

[19] Cunha, F., and J.J. Heckman, 2006, Identifying and estimating the distributions of ex post and ex ante returns to schooling: A survey of recent developments, mimeo.

[20] Cunha, F., J.J. Heckman, and S. Navarro, 2005, Separating uncertainty from heterogeneity in life cycle earnings, The 2004 Hicks Lecture, Oxford Econ. Papers, 57, 1-72.

[21] De la Fuente, A., 2003. Human capital in a global and knowledge-based economy. Part II: Assessment at the EU country level. Final Report, European Commission, Directorate General for Employment and Social Affairs.

[22] Duffie, D., and Y. Sun, 2007, Existence of independent random matching, The Annals of Applied Probability, 17, 386–419.

[23] Feldman, M., and C. Gilles, 1985, An expository note on individual risk without aggregate uncertainty, J. Econ. Theory, 35, 26-32.

[24] Flinn, C.J., and J. Mabli, 2008, Minimum wage and labor market outcomes under search with bargaining, IRP Discussion Paper 1337-08.

[25] Geanakoplos, J., and H. Polemarchakis, 1986, Existence, regularity and constrained suboptimality of competitive allocations when the asset market is incomplete, in W.P. Heller, R.M. Starr and D. Starrett (eds.), Uncertainty, Information and Communication: Essays in Honor of K.J. Arrow, Vol. III, Cambridge University Press, Cambridge, 65-95.

[26] Hosios, A.J., 1990, On the efficiency of matching and related models of search and unemployment, R. Econ. Studies, 57, 279-298.

[27] Judd, K.J., 1985, The law of large numbers with a continuum of i.i.d. random variables, J. Econ. Theory, 35, 19-25.

[28] Laing, D., T. Palivos, and P. Wang, 1995, Learning, matching and growth in a search model, R. Econ. Studies, 62, 115-129.

[29] Mendolicchio, C., D. Paolini, and T. Pietra, 2008, Human capital policies in a two-sector, static economy with imperfect markets, CRENoS Working Paper 2008-08.

[30] OECD, 2004, Education at a Glance, OECD, Paris.

[31] Petrongolo, B., C.A. Pissarides, 2001, Looking into the black box: A survey of the matching function, J. Econ. Lit., 39, 390-431.

[32] Pissarides, C.A., 2000, Equilibrium Unemployment Theory, MIT Press, II ed., Cambridge.

[33] Roy, A., 1951, Some thoughts on the distribution of earnings, Oxford Econ. Papers, 3, 135-146.

[34] Sattinger, M., 2003, A search version of the Roy model, mimeo.

[35] Shimer, R., 2005, The assignment of workers to jobs in an economy with coordination frictions, J. Pol. Econ., 2005, 113, 996-1025.

[36] Shimer, R., 2006, On-the-job search and strategic bargaining, Europ. Econ Review, 50, 811–830.

[37] Snyder, T.D., Dillow, S.A., and Hoffman, C.M., 2009, Digest of Education Statistics 2008 (NCES 2009-020), National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

[38] Sun, Y., 2006, The exact law of large numbers via Fubini extension and characterization of insurable risks, J. Econ. Theory, 126, 31-69.

[39] Sun, Y, and Y. Zhang, 2009, Individual risk and Lebesgue extension without aggregate uncertainty, J. Econ. Theory, 144, 432–443.

[40] Tawara, N., 2007, External returns to time-consuming schooling in a frictional labor market, mimeo, University of Chicago.

[41] Yashiv, E., 2003, Bargaining. the value of unemployment, and the behavior of aggregate wages, mimeo.

[42] Yashiv, E., 2006, Evaluating the performance of the search and matching model, Europ. Econ. Review, 50, 909–936.

Recently published

No.	Author(s)	Title	Date
14/2010	Stegmaier, J.	Effects of workplace representation on firm-provided further training in Germany	8/10
15/2010	Dauth, W.	The mysteries of the trade: Employment effects of urban interindustry spillovers	8/10
16/2010	Stüber, H. Beissinger, T.	Does downward nominal wage rigidity dampen wage increases?	8/10
17/2010	Schneider, J. Beblo, M.	Health an work – Indicators and determinnats: A revised literature and data review for Germany	9/10
18/2010	Konle-Seidl, R.	Make further vocational training pay: Possibilities and limits of promoting transitions by Public Employment Services (PES)	10/10
19/2010	Heineck, G. Wölfel, O.	Parental risk attitudes and children's secondary school track choice	10/10
20/2010	Dürnberger, A. Drasch , K. Matthes, B.	Kontextgestützte Abfrage in Retrospektiverhe-bungen: Ein kognitiver Pretest zu Erinnerungs-prozessen bei Weiterbildungsereignissen	10/10
21/2010	Hohmeyer, K. Wolff, J.	Direct job creation in Germany revisited: Is it effective for welfare recipients and does it matter whether participants receive a wage?	10/10
22/2010	Schlitte, Friso	Local human capital, segregation by skill, and skill-specific employment growth	11/10
1/2011	Grün, C. Mahringer, H. Rhein, Th.	Low-wage jobs: A means for employment inte-gration oft he unemployed? Evidence from ad-ministrative data in Germany and Austria	1/11
2/2011	Achatz, J. Trappmann, M.	Arbeitsmarktvermittelte Abgänge aus der Grund-sicherung: Der Einfluss von personen- und haushaltsgebundenen Arbeitsmarktbarrieren	2/11
3/2011	Patuelli, R. Schanne, N. Griffith, D.A. Nijkamp, P.	Persistence of regional unemployment: Applica-tion of a spatial filtering approach to local labour markets in Germany	2/11
4/2011	Riphahn, R.T. Trübswetter, P.	The intergenerational transmission of education-al attainment in East and West Germany	2/11
5/2011	Koller, L.	Lohnmobilität alleinstehender SGB-II-Leistungs-empfänger	2/11
6/2011	Brück-Kling berg, A. Burkert, C. Garloff, A. Seibert, H. Wapler, R.	Does higher education help immigrants find a job? A survival analysis	3/11
7/2011	Mendolicchio, C. Paolini, D. Pietra, T.	Income taxes, subsidies to education, and in-vestments in human capital	3/11

As per: 14.03.2011

For a full list, consult the IAB website
http://www.iab.de/de/publikationen/discussionpaper.aspx

Imprint

IAB-Discussion Paper 8/2011

Editorial address
Institute for Employment Research
of the Federal Employment Agency
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Regina Stoll, Jutta Palm-Nowak

Technical completion
Jutta Sebald

Website
http://www.iab.de

Download of this Discussion Paper
http://doku.iab.de/discussionpapers/2011/dp0811.pdf

For further inquiries contact the author:

Concetta Mendolicchio
Phone +49.911.179 2330
E-mail concetta.mendolicchio@iab.de