

City of dreams

Jorge De la Roca^{*‡}

University of Southern California

Gianmarco I. P. Ottaviano^{*¶}

*London School of Economics,
University of Bologna, CEP, and CEPR*

Diego Puga^{*§}

CEMFI and CEPR

April 2017

ABSTRACT: Higher ability workers benefit more from bigger cities while housing costs there are higher for everyone. However, there is little sorting on ability. We show this is partly because young individuals have an imperfect assessment of their ability, and, when they learn about it, early decisions have had a lasting impact and reduce their incentives to move. We formalize this idea through an overlapping generations model of urban sorting by workers with heterogeneous ability and self-confidence. Using data from the NLSY79, we find that the city-size choices of individuals vary with ability and self-confidence in line with our theoretical predictions.

Key words: cities, sorting, agglomeration, self-confidence, ability, learning

JEL classification: R10, R23

*We are grateful to Nathaniel Baum-Snow, Sanghoon Lee, and Jesse Shapiro for very helpful discussions and to Steve Gibbons, Pedro Mira, Theodore Papageorgiou, Frank Pisch, Steve Pischke, Anatoli Segura, Olmo Silva, and seminar participants at the LSE, CEMFI, and the NBER Summer Institute for comments and suggestions. Puga acknowledges funding from the European Research Council under the European Union's Seventh Research Framework Programme (ERC Advanced Grant agreement 269868 – SPYKES) and Horizon 2020 Programme (ERC Advanced Grant agreement 695107 – DYNURBAN) and from Spain's Ministerio de Economía y Competitividad (grant ECO2013-41755-P).

[‡]Sol Price School of Public Policy, University of Southern California, 650 Childs Way RGL 326, Los Angeles, CA 90089, USA (e-mail: jdelaroc@usc.edu; website: <http://jorgedelaroca.name>).

[¶]Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom (e-mail: g.i.ottaviano@lse.ac.uk; website: <http://scholar.google.com/citations?user=-z6U4wQAAAAJ&hl=en>).

[§]CEMFI, Casado del Alisal 5, 28014 Madrid, Spain (e-mail: diego.puga@cemfi; website: <http://diegopuga.org>).

1. Introduction

Bigger cities are associated with higher earnings and higher housing costs. Higher earnings take partly the form of a static premium that individuals obtain when they are working in a bigger city and partly the form of a dynamic gain due to accumulation of more valuable experience in bigger cities (Glaeser and Maré, 2001, Yankow, 2006, Baum-Snow and Pavan, 2012, De la Roca and Puga, 2017). These big-city benefits appear to be significantly larger for workers with higher ability within broad education or occupation categories (Baum-Snow and Pavan, 2012, De la Roca and Puga, 2017). Since more able workers benefit more from bigger cities, while housing costs there are higher for everyone regardless of ability, one might expect more able workers to locate disproportionately in bigger cities. And yet, this is not the case.

Several studies find that more educated workers or those with certain occupations are more likely to locate in bigger cities (Berry and Glaeser, 2005, Moretti, 2012, Davis and Dingel, 2013). However, within broad occupation or education groups, there appears to be little sorting on ability, whether this is measured through cognitive test results (Bacolod, Blum, and Strange, 2009), individual fixed-effects in a wage regression (De la Roca and Puga, 2017), measures of ability derived from a finite-mixture model in a structural estimation setting (Baum-Snow and Pavan, 2012), or individual residuals from a spatial equilibrium condition (Eeckhout, Pinheiro, and Schmidheiny, 2014).

A partial explanation for little sorting on ability is simply that many people are not mobile. In the United States, 56% of people live at age 40 in the same city or town where they were at age 14. Even for college-educated workers, the figure is 40%.¹ However, given that many people do move, one would expect them to take into account how they would fare in different cities depending on their ability. Our starting point in this paper is that it is not all that easy for individuals to make such a calculation. When young individuals choose a location, they may have a very imperfect assessment of their own ability and, by the time they learn enough about it, early decisions have had a lasting impact and reduce their incentives to move.

A large literature in psychology documents that individuals' assessment of their own ability generally has little resemblance to their actual ability (see Dunning, Heath, and Suls, 2004, for a survey). Correlation between people's views of their intelligence and their performance on intelligence tests and other academic tasks is typically between 0.2 and 0.3 (Hansford and Hattie, 1982). In the workplace, the correlation between how people expect to perform complex tasks and how they actually perform them is around 0.2 (Stajkovic and Luthans, 1998). Several complementary explanations for the pervasiveness of flawed self-assessment have been put forth. Not only assessing ability is inherently complex, but also assessing skills accurately often requires the same skills one is trying to assess (e.g., knowing whether one is good at maths requires sufficient mathematical knowledge). In addition, comparative assessments are very self-centred, relying largely on some loose perception of whether one is able to do something and not so much on how many others can do it substantially better. Even when people have information that would lead

¹These percentages are calculated from our panel data, described below.

them to more accurate self-assessments, they tend to neglect this information, which leads them to worse assessments than they are capable of (Benabou and Tirole, 2002).

We formalize the idea that flawed self-assessment can help explain limited sorting across cities of different sizes through an overlapping generations model. Workers in the model are heterogeneous in ability and in self-confidence, where the latter is defined as individuals' assessment of their own ability (in an extension, we also allow for heterogeneity in educational attainment). Relative to small cities, big cities allow young workers to gain more valuable experience and provide greater opportunities for more senior workers to apply their accumulated experience. Both advantages are stronger for high-ability workers. However, big cities also involve higher living costs. Young workers choose their location based on the benefits and costs of big cities and on their self-confidence, which may or may not correspond to their actual ability. They then accumulate experience depending on their chosen location, ability and luck. In the process, they also learn their own ability. Based on accumulated experience and ability, as well as on the opportunities and costs of big and small cities, senior workers choose whether to relocate or not.

The model predicts various patterns of bilateral sorting between big and small cities along workers' life cycles. Location decisions by young workers are mostly driven by self-confidence. For senior workers, ability plays a stronger role in determining location, but the lasting impact of earlier choices dampens their incentives to move. The imperfect self-assessment of young workers and the lasting consequences of early location choices combine to reduce the aggregate extent of sorting. Despite limited sorting, some workers who seriously underestimated their own ability relocate from a small to a big city once their labour market experience provides them with better information of their true capabilities. Workers who instead greatly overestimated their ability tend to fail in big cities and then move away. However, if they are fortunate enough to succeed they may stay in the big city to take advantage of greater local opportunities for experienced workers.

We test these and other predictions on panel data from the National Longitudinal Survey of Youth 1979 (NLSY79), which allows us to track individuals' location and labour market activities as well as a rich set of personal characteristics. Our primary measure of ability is the individual's percentile score in the Armed Forces Qualification Test (AFQT), a cognitive ability test that was administered to respondents when their median age was 19. In our model, we use the term self-confidence to refer to individuals' perception of their own ability. Prior to being provided their results on the AFQT, respondents in the NLSY79 were subject to a self-evaluation test devised by Rosenberg (1965), which has been found to measure well individuals' perception of their own ability to perform in a wide variety of tasks, in particular those that are job-related (Judge, Erez, and Bono, 1998, Chen, Gully, and Eden, 2001). We study individuals' location choices upon completing education (corresponding to the junior period of our model) and ten years later (the senior period).

When we examine the raw relationship between the location choices of individuals in their junior and senior periods and their levels of self-confidence and ability, we find that the data closely match our theoretical predictions. We then estimate logit models to look at the determinants of locating in a small or a big city either when junior or senior, while controlling for other drivers of mobility. Our findings confirm that individuals with higher levels of self-confidence are more likely to locate in a big city upon entering the job market. Instead, high-ability young individuals

are not significantly more likely to locate initially in a big city. When examining senior period relocations, we find that self-confidence no longer influences the decision to relocate from small to big cities whereas the level of ability is a significant determinant. While corrections to flawed self-assessment are a significant determinant of relocations from small to big cities, relocations from big to small cities appear to be driven instead by lack of success in the big city.

Our findings contribute to the literature on learning in cities. In particular, the model we develop has several elements in common with the model by Glaeser (1999): both are overlapping generations models where bigger cities facilitate learning. In our model, we introduce a second important difference between cities of different sizes: bigger cities also provide more opportunities to exploit previously acquired experience. More fundamentally, workers in Glaeser's (1999) model are homogeneous whereas workers in our model are heterogeneous in self-confidence and ability. This allows us to examine sorting patterns over the life cycle and the consequences of flawed self-assessment.

Thus, we also contribute to the literature on sorting across cities. Eeckhout, Pinheiro, and Schmidheiny (2014) and Davis and Dingel (2012) develop static models of sorting. In Eeckhout, Pinheiro, and Schmidheiny (2014) sorting is based on complementarities that are stronger between workers with extreme skill differences. They predict no sorting on average but a greater variance of skills in bigger cities. In Davis and Dingel (2012) there is perfect sorting driven by supermodularity. Behrens, Duranton, and Robert-Nicoud (2014) combine sorting, agglomeration and selection into a common theoretical framework. To keep the model manageable, they assume workers make an irreversible location choice, and obtain perfect sorting by heterogeneous ability, although variations in ex-post luck lead to heterogeneity in the productivity distribution. While in all of these models workers make a single location choice, in our dynamic model they choose their location in each period. Further, we introduce a role for workers' self-confidence so that the interplay between self-confidence, ability and luck shapes the incentives to relocate. Our analysis of the varying importance of ability and self-confidence over the workers' life cycle ties to earlier work on how the preferred urban environment can change, such as Duranton and Puga (2001), who study firms' as opposed to workers' life cycle.

Finally, our findings contribute to the literature on personality psychology and economics recently reviewed by Almlund, Duckworth, Heckman, and Kautz (2011), who document the power of personality traits both as predictors and as causes of academic and economic success, health, and criminal activity. For many outcomes, personality measures are just as predictive as cognitive measures derived from IQ and achievement tests, even after controlling for family background and cognition. Standard measures of cognition are also heavily influenced by personality traits which vary over the life cycle and can be altered by experience and investment. Urban economics has paid much attention to education and cognitive skills, but less so to other skills and personality traits. An exception is Bacolod, Blum, and Strange (2009), who show that workers with stronger cognitive and people skills (as inferred from their occupation and the skills associated with this occupation in the Dictionary of Occupational Titles) are more highly rewarded in bigger cities, while those with greater motor skills and physical strength are not. While it is possible that certain personality traits could also be more highly rewarded in big cities, in this paper we show that

this is not the case for self-confidence. Instead, self-confidence appears to matter for location decisions because it reflects individuals' perception of their own ability. In follow-up work, Bacolod, Blum, and Strange (2010) combine their same measures of skills inferred from occupations with test-based measures of skills and personality traits. They show there is little variation on the average prevalence of various skills and personality traits across cities of different sizes, although skills and personality traits do affect the propensity to migrate. While also looking at the role of skills and personality traits in determining location decisions, our focus is on the role of flawed self-assessment and the varying importance of ability and self-confidence in determining location over the workers' life cycle.

The rest of the paper is organized in five sections. Section 2 presents the model of sorting and learning in cities of different sizes. Section 3 solves for individual location choices taking relative city sizes as given. Section 4 solves for the general equilibrium endogenizing relative city sizes. Section 5 describes the data set that is used to obtain the empirical results presented in section 6. Section 7 concludes.

2. The model

Each worker lives for two periods. We refer to workers in the first period of their life as junior workers and to workers in the second period of their life as senior workers. In each of these two periods, each worker chooses whether to locate in a big city or in a small city. We use subscript B to denote big city variables and subscript S to denote small city variables. The sizes of cities are derived as an equilibrium outcome of the location decisions of all agents in section 4. In our baseline model, we do not consider differences in education; these are added as an extension in appendix C.

Workers have heterogeneous ability. All junior workers are engaged in a simple task and a worker's ability, denoted by α , is her actual probability of successfully completing this simple task. However, junior workers may have an inaccurate assessment of their own ability. We denote by σ self-confidence, defined as a junior worker's assessment of her own ability (i.e., her belief about α). While trying to complete their simple task workers learn about their true ability, so all senior workers know their α accurately.² to locate in big cities.

Junior workers who fail to complete their simple task get a low return, normalized to 0. Those who instead succeed at completing this simple task, get a high return $\pi_1 > 0$. In addition to a higher return, junior workers who successfully complete their simple task also gain experience that will be valuable when senior. The key advantage of locating in the big city for junior workers is that it allows them to accumulate more valuable experience, as suggested by Glaeser and Maré

²For simplicity, in the main text we derive results under the assumption that a junior worker is totally convinced that her ability is σ until she updates this belief to the actual value α in the process of completing the simple task. In appendix C we re-derive results allowing for uncertainty, so that junior workers realize their self-assessment may be inaccurate. Each worker receives an initial signal, on the basis of which she takes her ability to be σ in expectation, with variance θ . Under this alternative assumption, the specific thresholds determining the location strategies of junior workers change, but all our results remain qualitatively unchanged. The same appendix introduces differences in education. This addition does not change the main results either, but generates some sorting of junior workers by education as more educated workers tend

(2001), and consistent with the evidence presented in Baum-Snow and Pavan (2012) and De la Roca and Puga (2017). Successful junior workers in the big city gain experience e_B while successful junior workers in the small city gain experience e_S , where $0 < e_S < e_B < 1$. Junior workers who fail at completing their simple task gain zero experience.

Senior workers, at the very least, engage in a simple task. In addition, some senior workers are presented with an opportunity to also engage in a more complex task. Regarding the simple task, senior workers who already succeeded at this simple task as junior workers, can replicate what they did and complete their simple task with certainty. Senior workers who failed as junior workers can try again and succeed at their simple task when senior with probability given by their ability α . The simple tasks yields a return π_1 if successful and zero return if unsuccessful. Regarding the complex task, to complete this a senior worker must be faced with a relevant opportunity and also have prior experience from completing a simple task as a junior worker. The key advantage of locating in the big city for senior workers is that this offers them greater opportunities to exploit their previously acquired experience. Opportunities for engaging in a complex task arise with probability Ω_B in big cities compared with probability Ω_S in small cities, where $0 < \Omega_S < \Omega_B < 1$. If faced with a complex task, a senior worker's probability of success is αe , the product of her innate ability α and the experience acquired as a junior worker e , where $e = e_B$ if she completed a simple task in the big city, $e = e_S$ if she completed a simple task in the small city, and $e = 0$ if she failed to complete a simple task whatever her location was. Note that this implies a positive interaction between ability and the more valuable experience provided by big cities, again consistent with the evidence presented in De la Roca and Puga (2017).³ Restated, the higher the ability of an experienced worker the more she benefits from the greater opportunities provided by the big city. Successfully completing a complex task yields an extra return π_2 on top of π_1 .

The disadvantage of locating in the big city for both junior and senior workers is that it involves higher costs for housing and commuting, which we refer to as urban costs.⁴ These urban costs are γ_B in the big city and γ_S in the small city, with $0 < \gamma_S < \gamma_B$.⁵ Since each individual agent chooses her own location in each period taking city sizes as given, we initially treat γ_B and γ_S as parameters. In section 4 we go one step further by explicitly introducing commuting costs and a spatial housing market in a simple monocentric city model, which makes γ_B and γ_S a function of the (endogenous) population size of each city.

Anyone who failed to complete a simple task as a junior worker will choose to locate in the small city as a senior worker. This is because such a worker cannot benefit from the greater opportunities present in big cities ($\Omega_B > \Omega_S$), since tackling a complex task requires success at a simple task first.

³An alternative interpretation of the interaction αe is that the probability of success at completing a complex task equals the worker's experience independently of her ability, but the experience a junior worker acquires depends not only on her location but also on her ability, so that a worker with ability α accumulates experience αe_B when successful in the big city and experience αe_S when successful in the small city.

⁴The higher cost of living in big cities, mostly because of the higher costs of housing, is widely documented. Based on French data on house and land transactions and consumer expenditures, Combes, Duranton, and Gobillon (2016) estimate an elasticity of urban costs with respect to city population ranging from 0.04 to 0.10.

⁵Note that in the absence of any difference in urban cost ($\gamma_B = \gamma_S$), nobody would ever locate in S. Analogously, with no difference in the value of experience ($e_B = e_S$), nobody would locate in B when junior, while with no difference in opportunities ($\Omega_B = \Omega_S$), nobody would locate in B when senior.

For such a worker there is also no point in locating in the big city to acquire greater experience ($e_B > e_S$), since this would only be valuable in the future, and senior workers are in their final period. At the same time, the big city has the disadvantage of its higher urban costs ($\gamma_B > \gamma_S$).

Given that senior workers who were unsuccessful in their junior period locate in the small city, the expected utility attained by locating in city i as a junior worker and, conditional on earlier success, locating in city j as a senior worker is:

$$U_{ij}(\alpha) = -\gamma_i + \alpha\pi_1 + (1 - \alpha)(-\gamma_S + \alpha\pi_1) + \alpha(-\gamma_j + \pi_1 + \Omega_j \alpha e_i \pi_2), \quad i, j \in \{B, S\}. \quad (1)$$

Someone who locates in city $i \in \{B, S\}$ as a junior worker incurs an urban cost γ_i . With probability α she succeeds at completing the simple task in her junior period and obtains a return π_1 . In the senior period, if she failed at the simple task when junior (which happens with probability $1 - \alpha$), she then locates in S as a senior worker, incurring urban cost γ_S , and takes another shot at the simple task, which yields a return π_1 with probability α . If she instead succeeded at the simple task when junior (which happens with probability α), she locates in city $j \in \{B, S\}$ as a senior worker and incurs an urban cost γ_j . Then, she is guaranteed at least return π_1 , and with probability Ω_j she faces the opportunity to also engage in a more complex task. She successfully completes this complex task, yielding an additional return π_2 , with probability αe_i that depends on her ability and the experience e_i she acquired as a junior worker in city i . We can now use this equation to compare the possible location choices for each worker.⁶

3. Equilibrium location choices

A worker's location choice when junior will affect her location choice when senior. Thus, we must first examine the optimal location choice of senior workers conditional on their location when junior. Only then can we examine the optimal location choice of junior workers.

Senior period location

Consider a worker who locates in a big city when junior. She prefers to also locate in a big city when senior (conditional of earlier success) if and only if the expected utility from doing so, $U_{BB}(\alpha)$, is greater than the expected utility from relocating from a big to a small city, $U_{BS}(\alpha)$:

$$U_{BB}(\alpha) - U_{BS}(\alpha) = \alpha [\alpha(\Omega_B - \Omega_S)e_B \pi_2 - (\gamma_B - \gamma_S)] > 0, \quad (2)$$

Solving inequality (2) for α yields the equivalent condition

$$\alpha > \alpha_{BB>BS} \equiv \frac{\Delta\gamma}{e_B \pi_2 \Delta\Omega}, \quad (3)$$

where

$$\Delta\gamma \equiv \gamma_B - \gamma_S, \quad (4)$$

$$\Delta\Omega \equiv \Omega_B - \Omega_S. \quad (5)$$

⁶Equation (1) does not discount senior period variables. Introducing discounting by a factor δ would turn this equation into $U_{ij}(\alpha) = -\gamma_i + \alpha\pi_1 + \delta[(1 - \alpha)(-\gamma_S + \alpha\pi_1) + \alpha(-\gamma_j + \pi_1 + \Omega_j \alpha e_i \pi_2)]$. This would slightly complicate some of the expressions below but would not affect our results qualitatively.

The ability threshold defined by equation (3), $\alpha_{BB \succ BS}$, is such that anyone with ability above this threshold gets higher utility by locating in B as a junior worker and, conditional on success, also locating in B as a senior worker than by locating in B as a junior worker and relocating to S as a senior worker (hence the subscript $BB \succ BS$). We use this same notation for all thresholds that follow. Thus, junior workers who locate in the big city and successfully complete a simple task sort by ability when senior: those with high ability ($\alpha > \alpha_{BB \succ BS}$) stay in B , while those with low ability ($\alpha \leq \alpha_{BB \succ BS}$) relocate to S . Ability matters in the location choice of senior workers because they are willing to incur the high urban costs of the big city only in the hope of successfully completing a complex project and, other things equal, this is more likely the higher their ability. The ability threshold $\alpha_{BB \succ BS}$ is higher (fewer senior workers locate in B) the higher the urban cost gap between B and S ($\Delta\gamma$), the lower the difference in opportunities to engage in a complex task ($\Delta\Omega$), the lower the value of experience acquired as a successful junior worker in the big city (e_B), and the lower the extra return from completing a complex task (π_2).

Consider a worker who instead locates in S when junior. She prefers to relocate to B when senior (conditional of earlier success) if and only if

$$U_{SB}(\alpha) - U_{SS}(\alpha) = \alpha [\alpha(\Omega_B - \Omega_S)e_S \pi_2 - (\gamma_B - \gamma_S)] > 0, \quad (6)$$

or, equivalently, if and only if

$$\alpha > \alpha_{SB \succ SS} \equiv \frac{\Delta\gamma}{e_S \pi_2 \Delta\Omega}. \quad (7)$$

Thus, junior workers who locate in the small city and successfully complete a simple task also sort by ability when senior: those with high ability ($\alpha > \alpha_{SB \succ SS}$) relocate to B , while those with low ability ($\alpha \leq \alpha_{SB \succ SS}$) stay in S . The comparative statics for this threshold $\alpha_{SB \succ SS}$ are the same as for $\alpha_{BB \succ BS}$ except that, since we are now looking at the senior period decision of a worker who located in S in her junior period, it is e_S rather than e_B that appears in the threshold.

Note that the location choice as a junior worker affects the ability threshold above which a senior worker prefers to locate in the big city. A successful junior worker who locates in B acquires more valuable experience than one who locates in S ($e_B > e_S$), so she has a lower ability threshold above which locating in B when senior is worthwhile:

$$\alpha_{BB \succ BS} \equiv \frac{\Delta\gamma}{e_B \pi_2 \Delta\Omega} < \frac{\Delta\gamma}{e_S \pi_2 \Delta\Omega} \equiv \alpha_{SB \succ SS}. \quad (8)$$

Note also that the probability of not completing a simple task $1 - \alpha$ falls with ability but is nevertheless positive for everyone. Thus, even some very able junior workers fail and locate in the small city in their senior period. This is a first reason why sorting by ability will always be imperfect: ability is important but luck also plays a role.

Gathering all of the above, location decisions in the senior period follow the rules in the following Lemma.

Lemma 1. A worker who fails at the simple task when junior, locates in S when senior.

A worker with $\alpha \leq \alpha_{BB \succ BS}$ locates in S when senior.

A worker with $\alpha_{BB \succ BS} < \alpha \leq \alpha_{SB \succ SS}$ locates in B when senior if and only if she located in B when junior and succeeded at the simple task.

A worker with $\alpha_{SB>SS} < \alpha$ locates in B when senior unless she fails at the simple task when junior.

Junior period location when self-confidence accurately reflects ability

Having characterized the senior period location choice conditional on junior period location and success or failure at a simple task, we now turn to the junior period location choice. We begin with a simple case in which an individual's self-confidence while junior accurately reflects her ability ($\sigma = \alpha$). Afterwards, we will examine the more general case where an individual's self-confidence while junior may not reflect her actual ability ($\sigma \neq \alpha$); only after working in a first task does a worker learn her actual ability.

In Lemma 1 we have shown that senior-period location depends on the value of ability relative to two thresholds, $\alpha_{BB>BS}$ and $\alpha_{SB>SS}$, where $\alpha_{BB>BS} < \alpha_{SB>SS}$. Thus, in studying junior-period location, we will need to consider three ranges of ability: $\alpha \leq \alpha_{BB>BS}$, $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$, and $\alpha_{SB>SS} < \alpha$.

For a worker with low ability $\alpha \leq \alpha_{BB>BS}$ locating in B when senior is never worthwhile regardless of her junior period location. However, even knowing she will locate in S when senior, such a worker may nevertheless locate in B when junior to acquire more valuable experience. In particular, she locates in B in her junior period if and only if

$$U_{BS}(\alpha) - U_{SS}(\alpha) = \alpha^2 \Omega_S (e_B - e_S) \pi_2 - (\gamma_B - \gamma_S) > 0, \quad (9)$$

or, equivalently, if and only if

$$\alpha > \alpha_{BS>SS} \equiv \sqrt{\frac{\Delta\gamma}{\Omega_S \pi_2 \Delta e}}, \quad (10)$$

where

$$\Delta e \equiv e_B - e_S. \quad (11)$$

Ability matters in the location choice of junior workers for two reasons (hence the exponent of α in equation 9 and the square root in equation 10). First, because more able workers are more likely to complete a simple task and attain experience, which is more valuable if acquired in the big city. Second, because that experience helps complete a complex task once the worker reaches her senior period and success at that is also more likely the greater her ability. A worker with $\alpha \leq \alpha_{BB>BS}$, who always locates in S when senior, may be willing to incur the higher urban costs of B when junior in order to acquire additional experience, which could be valuable if an opportunity to use that experience in S when senior were to arise. The ability threshold $\alpha_{BS>SS}$ above which locating in B when junior is worthwhile for such a worker is higher (fewer junior workers locate in B) the higher the urban cost gap between B and S , $\Delta\gamma$, the lower the difference in the experience acquired by successful junior workers in B and S , Δe , the smaller the opportunities to engage in a complex task as a senior worker in the small city, Ω_S , and the lower the extra return from completing a complex task, π_2 . Note that these comparative statics are the same as for the ability thresholds that determine the location of senior workers, except that in the junior period it is the difference in experience, Δe , instead of the difference in opportunities, $\Delta\Omega$, that matters.

For a worker with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locating in B when senior is worthwhile if she located in B when junior and successfully completed her simple task. Knowing this, she locates in B in her junior period if and only if

$$U_{BB}(\alpha) - U_{SS}(\alpha) = \alpha^2(\Omega_B e_B - \Omega_S e_S) \pi_2 - (1 + \alpha)(\gamma_B - \gamma_S) > 0, \quad (12)$$

or, equivalently, if and only if

$$\alpha > \alpha_{BB>SS} \equiv \frac{1}{2} \left(\tilde{\alpha} + \sqrt{\tilde{\alpha}^2 + 4\tilde{\alpha}} \right), \quad \text{where} \quad \tilde{\alpha} \equiv \frac{\Delta\gamma}{(\Omega_B e_B - \Omega_S e_S) \pi_2}. \quad (13)$$

Note that the more complex functional form for this threshold $\alpha_{BB>SS}$ occurs because for workers with intermediate ability ($\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$) their junior period location affects their senior period location. Thus, in the utility comparison of equation (12) there is a difference in urban costs both for the junior period and, conditional on success, also for the senior period. In contrast, workers with low ability ($\alpha < \alpha_{BB>BS}$) chose their junior period location knowing they will locate in S when senior regardless. The comparative statics on this threshold $\alpha_{BB>SS}$ are nevertheless the same as for the threshold $\alpha_{BS>SS}$, except that for workers with intermediate ability the differences in opportunities combine with the differences in experience to determine their junior location.

For a worker with high ability $\alpha_{SB>SS} < \alpha$ locating in B when senior (conditional on junior period success) is always worthwhile regardless of her junior period location. Knowing she will locate in B when senior if successful, she locates in B in her junior period if and only if

$$U_{BB}(\alpha) - U_{SB}(\alpha) = \alpha^2 \Omega_B (e_B - e_S) \pi_2 - (\gamma_B - \gamma_S) > 0, \quad (14)$$

or, equivalently, if and only if

$$\alpha > \alpha_{BB>SB} \equiv \sqrt{\frac{\Delta\gamma}{\Omega_B \pi_2 \Delta e}}. \quad (15)$$

The comparative statics for this threshold $\alpha_{BB>SB}$ are the same as for $\alpha_{BS>SS}$ except that, since we are now looking at the junior period decision of a worker who will locate in B instead of in S when senior, it is Ω_B rather than Ω_S that appears in the threshold.

Gathering all of the above, location decisions in the two periods follow the rules in the following Lemma.

Lemma 2. A worker who fails at the simple task when junior, locates in S when senior.

A worker with $\alpha \leq \alpha_{BB>BS}$ locates in S when senior; if $\alpha \leq \alpha_{BS>SS}$ she also locates in S when junior while if instead $\alpha_{BS>SS} < \alpha$ she locates in B when junior.

A worker with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locates in S both periods if $\alpha \leq \alpha_{BB>SS}$; if instead $\alpha_{BB>SS} < \alpha$, she locates in B when junior and, if she succeeds at the simple task, also locates in B when senior.

A worker with $\alpha_{SB>SS} < \alpha$ locates in B when senior unless she fails at the simple task when junior; if $\alpha_{BB>SB} < \alpha$ she also locates in B when junior while if instead $\alpha \leq \alpha_{BB>SB}$ she locates in S when junior.

We now have all the information required to characterize location as a function of ability when workers' self-confidence while junior accurately reflects their ability ($\sigma = \alpha$).

Equilibrium location when self-confidence accurately reflects ability

At first glance, the positive interaction between workers' ability and the more valuable experience and greater opportunities that the big city provides would suggest a straightforward 'assortative matching' result with low-ability workers locating in S both periods and high-ability workers locating in B both periods, with the exception of those high-ability workers who, having failed at the simple task in B when junior, move to S when senior. The previous lemmas, however, hint at additional relocation trajectories. For instance, workers may decide to locate in S when junior to save on urban costs but then, if they are successful and acquire some experience, relocate to B when senior to take advantage of greater opportunities to put that experience to use. Alternatively, workers may decide to gain additional experience by locating in B when junior but then, even if they are successful, relocate to S when senior and save on urban costs. The existence of these trajectories clearly depends on whether the differences in experience or in opportunities between S and B dominate the tradeoffs that workers face given their ability.

The following proposition characterizes the exact conditions under which relocations may or may not take place in equilibrium and the direction of these relocations.

Proposition 1. When workers' self-confidence while junior accurately reflects their ability ($\sigma = \alpha$), location and relocation patterns fall in one of three cases.

Case 1. If $\frac{\Delta e}{\Delta \Omega^2} < \frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma}$:

- Workers with $\alpha \leq \alpha_{SB \succ SS}$ locate in S both periods.
- Workers with $\alpha_{SB \succ SS} < \alpha \leq \alpha_{BB \succ SB}$ locate in S when junior and, if and only if successful, relocate to B when senior.
- Workers with $\alpha_{BB \succ SB} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Case 2. If $\frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma} \leq \frac{\Delta e}{\Delta \Omega^2} \leq \frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma}$:

- Workers with $\alpha \leq \alpha_{BB \succ SS}$ locate in S both periods.
- Workers with $\alpha_{BB \succ SS} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Case 3. If $\frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma} < \frac{\Delta e}{\Delta \Omega^2}$:

- Workers with $\alpha \leq \alpha_{BS \succ SS}$ locate in S both periods.
- Workers with $\alpha_{BS \succ SS} < \alpha \leq \alpha_{BB \succ BS}$ locate in B when junior and relocate to S when senior.
- Workers with $\alpha_{BB \succ BS} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Proof See appendix A. □

Intuitively, in all three cases described by the proposition, workers with high ability locate in B both periods (provided they are successful at the simple task when junior) whereas workers with low ability locate in S both periods. This is explained by the positive interaction between workers' ability and the more valuable experience and opportunities provided by B . What differs across cases is the location of workers with intermediate ability.

In case 1, the difference between B and S in the value of experience acquired when junior is dominated by the difference in opportunities to use that experience when senior. Restated in terms of parameters, Δe is small relative to $\Delta\Omega$. This makes it worthwhile for workers of intermediate ability to locate in S when junior and in B when senior if successful at the simple task when junior. In this way, they avoid the higher urban costs of B when junior that, due to their moderate ability, are not compensated by a higher enough expected benefit in terms of differential experience. And yet, they relocate to B when senior since this provides sufficiently larger expected opportunities to offset the higher urban costs of B when senior.

In case 3, the situation is reversed: the difference between B and S in the value of experience acquired when junior dominates the difference in opportunities to use that experience when senior. Restated in terms of parameters, Δe is large relative to $\Delta\Omega$. This makes it worthwhile for workers of intermediate ability to locate in B when junior and in S when senior no matter whether they are successful or not at the simple task when junior. In this way they enjoy the more valuable experience associated with working in B when junior but avoid a higher urban cost when senior that, due to their moderate ability, is not compensated by a higher enough benefit in terms of differential opportunities.

In the intermediate case 2 neither the greater experience nor the greater opportunities of B dominate and changing location is not worthwhile irrespective of ability, even for workers who succeed at completing the simple task when junior.

The three cases are respectively described in panels (a), (b) and (c) of figure 1 along the 45-degree line. The diagram in each panel represents self-confidence (σ) on the horizontal axis and ability (α) on the vertical axis. Proposition 1 concerns the case where self-confidence accurately reflects ability ($\sigma = \alpha$), which corresponds to the diagonal. For instance, in panel (a) (for case 1) the segment of the diagonal in the bottom-left of the diagram ($\alpha \leq \alpha_{SB>SS}$) lies in an area marked SS , meaning that workers with low ability locate in S both periods. The intermediate segment of the diagonal ($\alpha_{SB>SS} < \alpha \leq \alpha_{BB>SB}$) lies in an area marked SB , meaning that workers with intermediate ability locate in S when junior and, if successful, relocate to B when senior. The top-right segment of the diagonal ($\alpha_{BB>SB} < \alpha$) lies in an area marked BB , meaning that workers with high ability locate in B both periods, unless they fail at the simple task when junior and relocate to S when senior. The area off the diagonal corresponds to cases where a worker's self-confidence when junior may not accurately reflect her actual ability, to which we turn next.

Equilibrium location when self-confidence does not reflect ability accurately

We now consider the situation in which a worker's self-confidence when junior may not accurately reflect her actual ability ($\sigma \neq \alpha$). Only after trying to complete a simple task for the first time does a worker realize her actual ability. To remain as general as possible, we do not make any

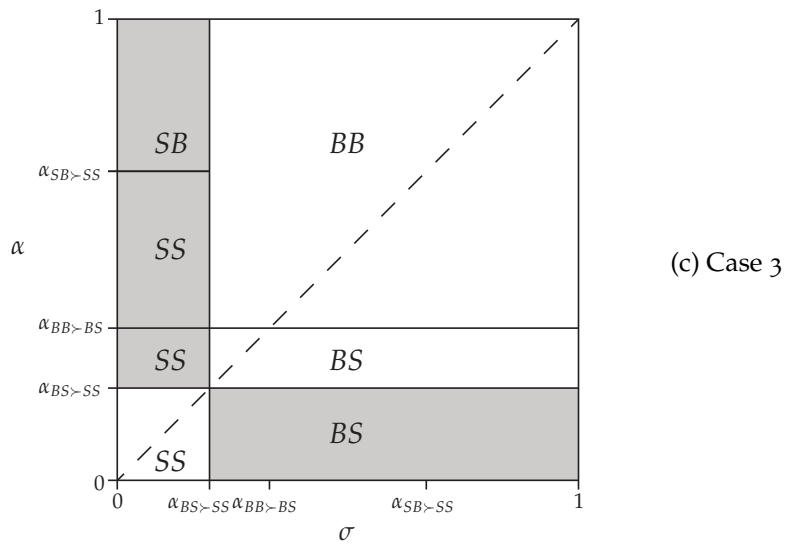
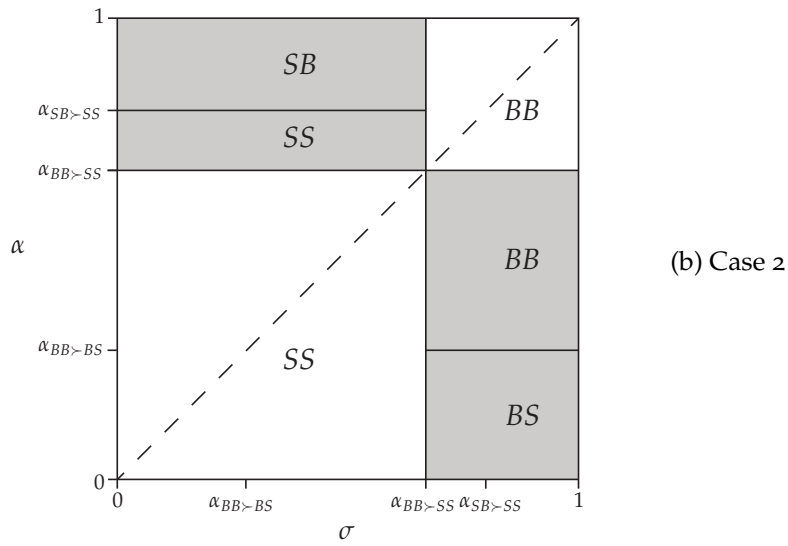
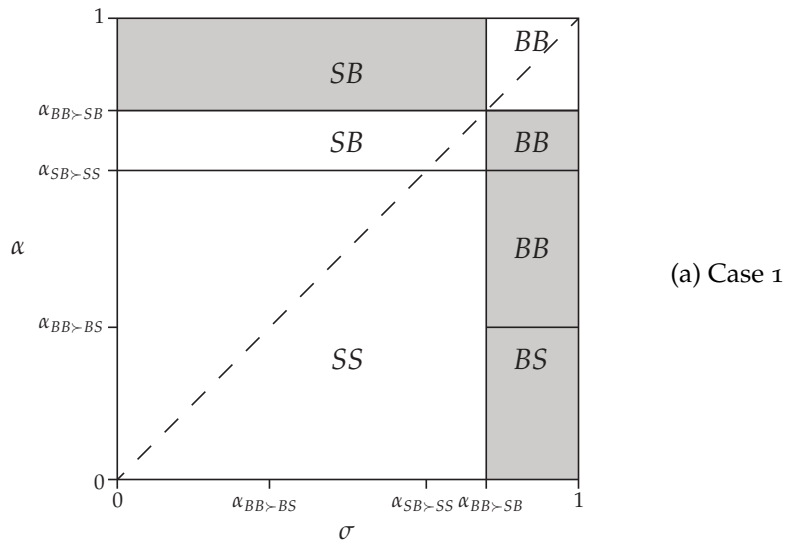


Figure 1: Equilibrium location choices by self-confidence and ability

specific assumption on the correlation between self-confidence and ability, characterizing instead the equilibrium location of workers for all possible combinations of σ and α .

When self-confidence in the junior period does not accurately reflect a worker's ability, both junior and senior location decisions change. The junior decision is still driven by the same tradeoffs as before but it is now based on the individual's self-confidence σ (i.e., her prior about her ability) rather than her ability α . The senior decision is also affected. While α is known to the senior worker, her junior period decision affects her experience, which in turn affects the relative incentives to locate in B or S when senior.

The main implication is that workers whose self-confidence is very different from their ability may end up making decisions they would not have made if they had known their actual ability to start with. The resulting patterns of self-deceit are summarized in the following proposition.

Proposition 2. When workers' self-confidence while junior does not reflect ability accurately ($\sigma \neq \alpha$), location and relocation patterns fall in one of three cases.

Case 1. If $\frac{\Delta e}{\Delta \Omega^2} < \frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma}$:

- During their junior period
 - Workers with $\sigma \leq \alpha_{BB \succ SB}$ locate in S .
 - Workers with $\alpha_{BB \succ SB} < \sigma$ locate in B .
- During their senior period
 - Workers with $\alpha \leq \alpha_{BB \succ BS}$ locate in S .
 - Workers with $\alpha_{BB \succ BS} < \alpha \leq \alpha_{SB \succ SS}$ locate in B if $\alpha_{BB \succ SB} < \sigma$ and they succeed at the simple task when junior; they locate in S otherwise.
 - Workers with $\alpha_{SB \succ SS} < \alpha$ locate in B if they succeed at the simple task when junior; they locate in S otherwise.

Case 2. If $\frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma} \leq \frac{\Delta e}{\Delta \Omega^2} \leq \frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma}$:

- During their junior period
 - Workers with $\sigma \leq \alpha_{BB \succ SS}$ locate in S .
 - Workers with $\alpha_{BB \succ SS} < \sigma$ locate in B .
- During their senior period
 - Workers with $\alpha \leq \alpha_{BB \succ BS}$ locate in S .
 - Workers with $\alpha_{BB \succ BS} < \alpha \leq \alpha_{SB \succ SS}$ locate in B if $\alpha_{BB \succ SS} < \sigma$ and they succeed at the simple task when junior; they locate in S otherwise.
 - Workers with $\alpha_{SB \succ SS} < \alpha$ locate in B if they succeed at the simple task when junior; they locate in S otherwise.

Case 3. If $\frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma} < \frac{\Delta e}{\Delta \Omega^2}$:

- During their junior period

- Workers with $\sigma \leq \alpha_{BS>SS}$ locate in S .
- Workers with $\alpha_{BS>SS} < \sigma$ locate in B .
- During their senior period
 - Workers with $\alpha \leq \alpha_{BB>BS}$ locate in S .
 - Workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B if $\alpha_{BS>SS} < \sigma$ and they succeed at the simple task when junior; they locate in S otherwise.
 - Workers with $\alpha_{SB>SS} < \alpha$ locate in B if they succeed at the simple task when junior; they locate in S otherwise.

Proof Location decisions when junior follow directly from Proposition 1, but with ability α replaced by self-confidence σ .

Locations decisions when senior depend on the value of α relative to the thresholds $\alpha_{BB>BS}$ and $\alpha_{SB>SS}$.

Starting from the lowest ability, consider first workers with $\alpha \leq \alpha_{BB>BS}$. By Lemma 1, such workers always locate in S when senior.

Consider next workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$. By Lemma 1, workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B when senior if and only if they locate in B when junior and succeed at the simple task. We now need to distinguish the same three cases as in Proposition 1. In case 1, by Proposition 1 (with α replaced by σ), workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B when junior (and thus also when senior if successful at the simple task) if and only if $\alpha_{BB>SB} < \sigma$. In case 2, by Proposition 1 (with α replaced by σ), workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B when junior (and thus also when senior if successful at the simple task) if and only if $\alpha_{BB>SS} < \sigma$. In case 3, by Proposition 1 (with α replaced by σ), workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B when junior (and thus also when senior if successful at the simple task) if and only if $\alpha_{BS>SS} < \sigma$.

Turning finally to workers with $\alpha_{SB>SS} < \alpha$, by Lemma 1, such workers always locate in B when senior if they succeed at the simple task when junior; they locate in S otherwise. \square

The three cases are again described in the three panels of figure 1, the focus being now on the area away from the 45-degree line where self-confidence and ability do not coincide ($\sigma \neq \alpha$). The intuition behind the existence of these three cases is the same as before: in the first case the difference in opportunities between B and S dominates; in the second case neither the difference in experience nor the difference in opportunities prevails; in the third case it is the difference in experience that dominates. The novelty is the presence of new trajectories that did not arise before and correspond to the shaded areas in the three figures.

In panel (a) of figure 1, for case 1, where the difference in opportunities between B and S dominates, most new trajectories arise for workers who locate in B when junior due to their overconfidence, while had they accurately assessed their ability when young, they would have started in S instead. Their ranges of ability and self-confidence ($\alpha < \alpha_{BB>SB} < \sigma$) correspond to the three shaded rectangles on the right of panel (a). What happens to them depends on how high their ability turns out to be within this range.

A first group consists of workers with high-enough ability to locate in B when senior no matter where they have completed the simple task when junior ($\alpha_{SB>SS} < \alpha < \alpha_{BB>SB}$). Their ability, however, would not be high enough to justify location in B when junior if correctly assessed. Yet, their overconfidence brings them to B when junior and they remain there if they are lucky enough to complete the simple task. These are workers whose σ and α fall in the top right shaded rectangle labelled BB .

A second group consists of workers with slightly lower ability ($\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$) who also locate in B when junior due to overconfidence (and are lucky enough to gain experience e_B). These are workers whose σ and α fall in the middle right shaded rectangle labelled BB . Going horizontally from this rectangle to the diagonal, we see that these workers, had they accurately assessed their ability when junior, would have located in S both periods. However, given that they located in B when junior driven by their overconfidence and gained some valuable experience, and in light of their intermediate ability, now it is worthwhile for them to remain in B when senior.

A third group consists of workers who decide to locate in B when junior unaware of their very low ability ($\alpha \leq \alpha_{BB>BS}$). Some of them are lucky: they succeed at the simple task and gain higher experience e_B . However, while completing the simple task, they realize that their ability is too low to stand a good enough chance of exploiting higher opportunities Ω_B by remaining in B , so they relocate to S when senior. This holds for all workers whose σ and α fall in the bottom right shaded rectangle labelled BS . Going horizontally from this rectangle to the diagonal, we see that these workers, had they accurately assessed their ability when young, would have instead located in S both periods.

In addition to the three new trajectories for overconfident workers, in panel (a) there is also one new trajectory for underconfident workers. Underconfident workers with very high ability locate in S when junior and, if successful as junior, move to B when senior once they realize that their ability is high enough to exploit better opportunities there. These are workers whose σ and α fall in the top left shaded rectangle labelled SB .

Panel (b) of figure 1 corresponds to case 2, where neither the difference in opportunities nor the difference in the value of experience between B and S dominate. The location of junior workers is again determined by their self-confidence, with the $\alpha_{BB>SS}$ being in this case the self-confidence threshold above which workers initially locate in B . The location choices of these workers when senior depend on the combination of their initial location choice and their actual ability.

Overconfident workers with very low ability locate in B when junior and, even if successful as junior, relocate to S when senior once they realize that their ability is too low to benefit from better opportunities in B . These are workers whose σ and α fall in the bottom right shaded rectangle labelled BS . Then there are other overconfident workers of higher ability who when junior are brought to B by their overconfidence and are lucky enough to succeed at the simple task. Thanks to the higher experience gained and their higher ability, it is beneficial for them to remain in B when senior. These are workers whose σ and α fall in the middle right shaded rectangle labelled BB . The proportions of overconfident workers who stick with their initial decision to locate in B and who instead prefer to relocate to S depend on the magnitude of the threshold $\alpha_{BB>BS}$. If $\Delta\Omega$ is large so that big cities provide substantially more opportunities to use previously acquired

experience, then most workers who initially locate in B will stay there even if their ability is low, provided they succeed at the simple task when junior.

Conversely, underconfident workers with very high ability locate in S when junior and, if successful as junior, move to B when senior once they realize that their ability is high enough to exploit better opportunities there. These are workers whose σ and α fall in the top left shaded rectangle labelled SB . Then there are other underconfident workers of lower ability who when junior are brought to S by their underconfidence. Even if they succeed at the simple task, doing so in S will lead them to accumulate less valuable experience than in B . As a result, having located in S initially leads them to stay there when senior whereas had they known their true ability they would have located in B both periods. These are workers whose σ and α fall in the middle left shaded rectangle labelled SS . The proportions of underconfident workers who stick with their initial decision to locate in S and who instead prefer to relocate to B depend on the magnitude of the threshold $\alpha_{SB>SS}$. Again, if $\Delta\Omega$ is large so that big cities provide substantially more opportunities to use previously acquired experience, then most workers who initially locate in S will nevertheless relocate to B , provided they succeed at the simple task when junior.

Panel (c) at the bottom of figure 1 corresponds to case 3, where the difference in the value of experience between B and S dominates. In this case, most new trajectories are driven by workers who locate in S when junior due to their underconfidence ($\sigma < \alpha$). Believing it is not worthwhile for them to incur the higher cost of B , they locate in S when junior while, had they accurately assessed their ability when young, they would have located in B instead. What happens to them depends on how high their ability turns out to be.

Underconfident workers with very high ability ($\alpha_{SB>SS} < \alpha$), as long as they succeed at completing the simple task when junior, find it worthwhile moving to B when senior. Thus, their senior location is not affected, although, having gained less experience by locating in S when junior, they are nevertheless less likely to succeed at the complex task than if they had located in B . These are workers whose σ and α fall in the top left shaded rectangle labelled SB .

For those among underconfident workers with intermediate ability ($\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$) locating in S when junior affects their senior location choice. Had they accurately assessed their ability when young, they would have located in B both periods. Having instead located in S when junior due to underconfidence, they gain less valuable experience even if successful and now do not find it worthwhile to locate in B when senior given that the difference in opportunities between B and S is limited. These are workers whose σ and α fall in the higher of the two left shaded rectangles labelled SS .

Underconfident workers with lower ability ($\alpha_{BS>SS} < \alpha \leq \alpha_{BB>BS}$) would not have located in B when senior anyway. Their underconfidence merely leads them to locate in S instead of B when junior. These are workers whose σ and α fall in the lower of the two left shaded rectangles labelled SS .

Finally, some overconfident workers with very low ability ($\alpha \leq \alpha_{BS>SS} < \sigma$) follow the same pattern as in case 1. Driven by their overconfidence, they locate in B when junior. Then, realizing that their ability is too low, they move to S when senior even if successful at the simple task. Had they correctly anticipated their low ability, they would have located in S both periods. This holds

for all workers whose σ and α fall in the bottom left shaded rectangle labelled BS .

Despite the differences among the three cases we have just described, they all have much in common. The location of junior workers is determined by their self-confidence, although the self-confidence threshold above which workers initially locate in B is different in each case. In all three cases, among workers who were underconfident and located in S , when locating in B would have been preferable, those with the highest ability relocate to B when senior while those with lower ability stick with their initial choice. Similarly, among workers who were overconfident and located in B , when locating in S would have been preferable, those with the lowest ability relocate to S when senior while those with higher ability stick with their initial choice. The main difference is that in case 1 and case 3 some junior workers who have an accurate initial assessment of their intermediate ability, and who succeed at the simple task, relocate between the junior and the senior period, in case 1 from S to B and in case 3 from B to S . In case 2 all relocations are driven either by flawed self-assessment or failure at the simple task when junior.

4. Endogenizing urban structure

We have so far determined individual location choices taking as given that there are big and small cities. However, city sizes in fact result from the combination of the location choices of individuals, and we must make sure there is consistency between individual location choices and city sizes. In other words, we must make sure that the equilibrium we have characterised exists. For this reason, we now endogenize the urban structure and solve for the general equilibrium of our model. Suppose each of the two cities B and S is linear and monocentric.⁷ Land covered by each city is endogenously determined and can be represented by a segment on the positive real line. All workers in a city perform their job at a single point $x = 0$, the Central Business District (CBD).

Workers consume housing and a freely tradable numéraire good. For simplicity, let us assume that all residences have the same size, are built under perfect competition with a constant capital to land ratio, and are owned by absentee landlords.⁸ Thus, every individual consumes one unit of floorspace built on one unit of land with a fixed amount of capital. The price of capital is constant throughout the economy while the price of land varies. Commuting costs increase linearly with distance to the CBD, so that a worker living at distance x incurs a commuting cost τx . The total urban costs for a worker located in a residence at a distance x from the CBD of city i are the sum of her commuting costs τx and her housing costs $P_i(x)$:

$$\gamma_i(x) = \tau x + P_i(x), \quad i, j \in \{B, S\}. \quad (16)$$

As a result, any resident in a city is willing to bid τx more for a house that is x closer to the CBD. Equilibrium house prices are then such that the increase in commuting costs incurred as one relocates towards the CBD is exactly offset by an increase in house prices.

⁷We develop a highly simplified version of the monocentric city model (Alonso, 1964, Mills, 1967, Muth, 1969). For an exposition of more general versions of the monocentric city model, see Brueckner (1987) and Duranton and Puga (2015).

⁸Having instead common ownership of the housing stock by local residents yields essentially the same results. One simply gets $\gamma_i = \frac{1}{2}\tau N_i$ instead of $\gamma_i = \tau N_i$ in equation (18) below.

Using N_i to denote the equilibrium population in city i , house prices in city i can then be expressed as

$$P_i(x) = \tau(N_i - x) + \bar{r}, \quad i, j \in \{B, S\}, \quad (17)$$

where the constant \bar{r} is the sum of the rental cost of the fixed amount of capital used in every residence and the rental price of land in the best non-urban use (e.g., agriculture). A worker living at the edge of a city has to commute a distance equal to the population of the city, thus incurring a commuting cost τN_i , but only pays \bar{r} for housing. A worker living at the CBD has no cost of commuting but pays an additional τN_i for her house. Substituting equation (17) into (16) yields urban cost in city i :

$$\gamma_i = \tau N_i + \bar{r}, \quad i, j \in \{B, S\}. \quad (18)$$

In order to allow for the coexistence of junior and senior workers in a city, let us assume that there are overlapping generations of workers. Each generation is made up of a continuum of workers of measure 1 and lives for two periods. Thus, workers coexist when junior with senior workers of the previous generation and coexist when senior with junior workers of the next generation. Since our focus is on the steady state, we avoid using a time subscript for our variables.

The total population of city i , N_i , is the sum of junior and senior workers in the city. Let us denote by n the difference in population between the big and the small city:

$$n \equiv N_B - N_S. \quad (19)$$

Note that $0 \leq n \leq 2$ since, by definition, the big city has a larger population and since the total population in the economy at any point in time is made up of two living generations with unit population mass each. Combining equations (18) and (19), we can then express the difference in urban costs between B and S , $\Delta\gamma \equiv \gamma_B - \gamma_S$, as

$$\Delta\gamma = \tau n. \quad (20)$$

Taking n as given, each worker can calculate $\Delta\gamma$ as per equation (20). She can then substitute this into equation (3) to calculate $\alpha_{BB \succ BS}$, into (7) to calculate $\alpha_{SB \succ SS}$, into (10) to calculate $\alpha_{BS \succ SS}$, into (13) to calculate $\alpha_{BB \succ SS}$, and into (15) to calculate $\alpha_{BB \succ SB}$. Given all these thresholds, each worker chooses her optimal location as per proposition 2. If we then add up how many workers choose to locate in each city, an equilibrium arises when this yields a difference in population between the two cities equal to n .

Adding $N_B + N_S = 2$ to equation (19) and solving for N_B , we can express population in B in terms of n :

$$N_B = 1 + \frac{n}{2}. \quad (21)$$

In equilibrium this must equal the total number of junior and senior workers choosing to reside in B , which we will denote by $b(n)$. To obtain an expression for $b(n)$, we must refer back to proposition 2. Proposition 2 distinguishes three cases depending on the value of $\Delta\gamma = \tau n$. Expressing the corresponding conditions in terms of n , case 1 arises for $0 \leq n < \underline{n}$, where $\underline{n} \equiv \frac{\pi_2 e_B^2 \Delta \Omega^2}{\tau \Omega_B \Delta e}$; case 2 arises for $\underline{n} \leq n \leq \bar{n}$, where $\bar{n} \equiv \frac{\pi_2 e_B^2 \Delta \Omega^2}{\tau \Omega_S \Delta e}$; and case 3 arises for $\bar{n} < n \leq 2$. Let us denote by

$f(\sigma, \alpha)$ the probability density function for the bivariate distribution of ability and self-confidence for workers in the population. Hence, we can write:

$$b(n) = \begin{cases} \int_{\alpha_{BB>SB}(n)}^1 \int_0^1 f(\sigma, \alpha) d\alpha d\sigma + \int_{\alpha_{BB>SB}(n)}^1 \int_{\alpha_{BB>BS}(n)}^{\alpha_{SB>SS}(n)} \alpha f(\sigma, \alpha) d\alpha d\sigma + \int_0^1 \int_{\alpha_{SB>SS}(n)}^1 \alpha f(\sigma, \alpha) d\alpha d\sigma & \text{if } 0 \leq n < \underline{n} \\ \int_{\alpha_{BB>SS}(n)}^1 \int_0^1 f(\sigma, \alpha) d\alpha d\sigma + \int_{\alpha_{BB>SS}(n)}^1 \int_{\alpha_{BB>BS}(n)}^{\alpha_{SB>SS}(n)} \alpha f(\sigma, \alpha) d\alpha d\sigma + \int_0^1 \int_{\alpha_{SB>SS}(n)}^1 \alpha f(\sigma, \alpha) d\alpha d\sigma & \text{if } \underline{n} \leq n \leq \bar{n} \\ \int_{\alpha_{BS>SS}(n)}^1 \int_0^1 f(\sigma, \alpha) d\alpha d\sigma + \int_{\alpha_{BS>SS}(n)}^1 \int_{\alpha_{BB>BS}(n)}^{\alpha_{SB>SS}(n)} \alpha f(\sigma, \alpha) d\alpha d\sigma + \int_0^1 \int_{\alpha_{SB>SS}(n)}^1 \alpha f(\sigma, \alpha) d\alpha d\sigma & \text{if } \bar{n} < n \leq 2 \end{cases} \quad (22)$$

Equation (22) can be readily understood by referring back to proposition 2. For example, the first case (for $0 \leq n < \underline{n}$) has three types of workers choosing to locate in B (each type captured by one of the three double integrals for this first case): junior workers with high self-confidence $\alpha_{BB>SB} < \sigma$; senior workers with intermediate ability $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ who in their junior period located in B due to high-self-confidence $\alpha_{BB>SB} < \sigma$ and succeeded at the simple task (probability α); and senior workers with high ability $\alpha_{SB>SS} < \alpha$, regardless of their self confidence, provided they succeeded at the simple task when junior (probability α).

We can also interpret equation (22) in terms of figure 1. Given the unit population mass of each generation of workers, the number of junior workers who decide to reside in B is given by the fraction of them with self-confidence and ability in rectangles BB or BS . The number of senior workers who decide to reside in B is given by the fraction of them with self-confidence and ability in rectangles BB or SB weighted by the probability α that they successfully completed the simple task when junior.

Any equilibrium value of n has to satisfy $b(n) = 1 + \frac{n}{2}$ for $0 \leq n \leq 2$. Under the assumption that $f(\sigma, \alpha)$ is continuous and differentiable in $\alpha \in [0, 1]$ and $\sigma \in [0, 1]$, the following result holds.

Proposition 3. There exists a unique equilibrium allocation of population across cities. In equilibrium, both the big and small cities are populated. The difference n in population between the big and small cities decreases with the common commuting cost per unit of distance τ , and increases with the additional opportunities $\Delta\Omega$ and the additional experience Δe provided by the bigger city.

Proof See appendix B. □

When deciding whether to locate in B , junior workers trade off the greater experience they are likely to acquire by locating there against the higher urban costs they need to incur. Senior workers trade off the greater opportunities B provides to use their previously-acquired experience against its higher urban costs. In equilibrium, some workers strictly prefer to locate in B and others strictly prefer to locate in S . Individual choices depend on self-confidence, ability and luck, all of

which vary across workers, on common parameters capturing the magnitude of the advantages and disadvantages of locating in the big city, and on the choices of all other workers.

In equilibrium, the difference in population n between B and S is such that the difference between the mass of workers who prefer to locate in B and the mass of workers who prefer to locate in S aggregates up to precisely n . Off-equilibrium, the mass of workers who given n prefer B to S may aggregate up to more than n , but then as more workers locate in B and fewer in S commuting and housing costs increase in B relative to S until an equilibrium is restored. And conversely, the reverse adjustment occurs if the mass of workers who given n prefer B to S aggregates up to less than n .

The comparative statics for equilibrium differences in city sizes are fairly intuitive. A higher cost of commuting per unit of distance (τ) implies a larger gap in urban costs for any given difference in population between B and S , and so results in a smaller equilibrium difference in population sizes (n). The greater the additional opportunities ($\Delta\Omega$) and the additional experience (Δe) provided by B , the more attractive is B relative to S , so a higher difference in population (n) is needed to balance things out in equilibrium.

5. Data

We use panel data from the “cross-sectional sample” of the National Longitudinal Survey of Youth 1979 (NLSY79). The survey, conducted by the US Department of Labor’s Bureau of Labor Statistics, follows a nationally representative sample of 6,111 men and women who were 14–22 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994 and were interviewed on a biennial basis since 1996. The NLSY79 contains information on a rich set of personal characteristics and tracks individuals’ labour market activities.

Locations

The confidential geocoded portion of the NLSY79 gathers information on the location of each respondent at multiple points in time. Specifically, for each respondent we know the county and state where they are located at birth, at age 14, and at each interview date since 1979. We use this location information to link the counties of location of each respondent to Core Based Statistical Areas (CBSA) as defined in 2008. A CBSA or metropolitan area is a collection of counties that delimits a local labour market.⁹ We classify individuals as located in a big city if they are within a CBSA with a population over two million in 2010. By this definition, 44 CBSA metropolitan areas are classified as big (from Kansas City with a population of just over 2 million to New York with almost 19 million). This is in line with other papers dealing with urban sorting, which typically classify cities as big when their population is above a threshold of between 1.5 million (Baum-Snow and Pavan, 2012) and 2.5 million (Eeckhout, Pinheiro, and Schmidheiny, 2014). Our results are very similar using these alternative definitions.

⁹ Core Based Statistical Areas (CBSA) are defined by the Office of Management and Budget. These CBSA metropolitan areas have replaced the metropolitan areas that were defined based on the 1990 census.

Ability

Importantly for us, the NLSY79 contains test results that aim to capture cognitive ability as well as self-evaluation. Our basic measure of ability is the individual's percentile score in the Armed Forces Qualification Test (AFQT), a cognitive ability test that was administered to NLSY79 respondents in 1980, when their median age was 19. Note that, while AFQT scores are used as a criterion of enlistment eligibility by the US military, the test was administered to all NLSY79 respondents regardless of whether they had any interest in the military in order to update norms for the test based on a nationally representative sample of young people.

Self-confidence

In our model, we use the term self-confidence to refer to individuals' perception of their own ability. Psychologists often use the term 'general self-efficacy' to capture this aspect of self-evaluation. This is defined as "individuals' perception of their ability to perform across a variety of different situations" (Judge, Erez, and Bono, 1998, p. 170). While the NLSY79 does not measure general self-efficacy per se, it does measure self-esteem, which is strongly related to it. Self-esteem is defined as "the overall value one places on oneself as a person" (Harter, 1990, p. 67). Conceptually, general self-efficacy and self-esteem are somewhat different aspects of self-evaluation in that self-esteem is a broader concept. However, there is a very strong empirical association between them. Respondents of the NLSY79 were subject in 1980 to a test to measure their self-esteem using Rosenberg's (1965) self-esteem scale. Summarising extant results on the relationship between Rosenberg's measure and general self-efficacy, Chen, Gully, and Eden (2001, p. 67) note that "the standard general self-efficacy scale is correlated highly with the Rosenberg (1965) self-esteem scale ($r = .75$ to $.91$)" and conclude that general self-efficacy "does not capture a construct distinct from self-esteem." Judge, Erez, and Bono (1998) argue that both concepts are strongly related to individuals' assessment of their own ability to perform on the job.

The Rosenberg (1965) measure is based on a ten-item questionnaire that assesses the self-evaluation of respondents through their expressed agreement or disagreement with various statements (e.g., "I am able to do things as well as most other people"). Five of the items have positively-worded statements and are assigned a score between 0 and 3 based on increasing agreement with the statement (strongly disagree, disagree, agree, strongly agree). Five of the items have negatively-worded statements and are assigned a score between 0 and 3 based on increasing disagreement with the statement. The Rosenberg measure is calculated by adding up the scores for all ten items. We convert the measure into a percentile score and use that as our measure of self-confidence. For us, it is important that this measure predates labour market experience since labour market outcomes could feed back into self-confidence (see the discussion in section 8 of Almlund, Duckworth, Heckman, and Kautz, 2011). It is also worth noting that respondents were subject to the Rosenberg test before they were given their results on the AFQT test of cognitive ability. The low correlation between the AFQT and Rosenberg scores, 0.3, reflects the pervasiveness of flawed self-evaluation and is in line with that reported in psychology studies

such as the aforementioned Hansford and Hattie (1982).¹⁰

Periods

To test the implications of our model, we need to define two periods (junior and senior) and relate respondents' location trajectories to their levels of ability and self-confidence. To match our model, we would like to use as the junior period for each individual the time immediately prior to entering the labour market and to use as the senior period a time that is late enough that he or she has accumulated significant labour market experience.

The NLSY79 records detailed information on the educational attainment of respondents over time, so that in each wave, we know their highest grade completed and their schooling enrolment status. We set the junior period for all respondents at the year after the highest level of education is completed.¹¹ The median age of individuals in their junior period is 20 for individuals with no post-secondary education and 23 for individuals with post-secondary education. We then determine whether each individual was located in a big metropolitan area or not in this junior period.

Next, we set the senior period for all respondents by adding ten years to their junior period. Thus, the median age of individuals in their senior period is 30 for individuals with no post-secondary education and 33 for individuals with post-secondary education. Again, we determine whether each individual was located in a big metropolitan area or not in this senior period.

Our starting sample is made of 6,111 individuals. We exclude individuals for whom the AFQT or the Rosenberg self-esteem scores are missing, which reduces the sample to 5,622 individuals. We are able to determine the junior period location of 5,413 individuals. Availability of the demographic controls we use further reduces our sample to 5,255.¹²

6. Empirical evidence

Observed location choices by self-confidence and ability

We begin by examining how the location choices of individuals vary with self-confidence and ability. To better illustrate location choices graphically and to relate these choices to the theoretical predictions depicted in figure 1, we first divide both the self-confidence and the ability measures into terciles. This yields nine possible combinations of self-confidence and ability terciles. Figure 2

¹⁰While there is sometimes the presumption that women are on average less self-confident than men, in our data Rosenberg's (1965) self-esteem scores are almost identical for men and women in mean and standard deviation. The correlation between the AFQT and Rosenberg scores is also 0.3 for both men and women.

¹¹We exclude educational periods that take place after a spell of more than two years away from education. For example, if an individual completes an undergraduate university degree, works for three or more years, and then goes back to university to pursue postgraduate studies, we take the year after completing the undergraduate degree as this individual's junior period, not the year after completing the postgraduate degree. In addition, for individuals who complete their highest level of education before turning 18, we use the year in which they turn 18 as their junior period. We exclude individuals who are older than 30 when they complete their highest level of education without any gap.

¹²Since the NLSY79 became biennial after 1994, for some individuals there is no interview ten years after their junior period and we must use the preceding or subsequent year.

plots in a grid each of those nine combinations of self-confidence and ability, with self-confidence on the horizontal axis and ability on the vertical axis.

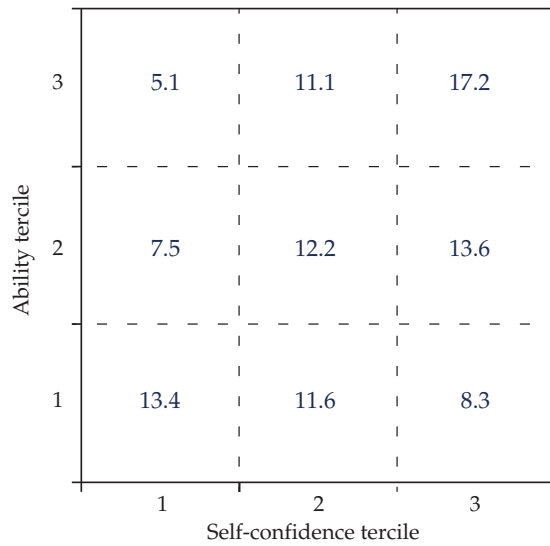
In panel (a) of figure 2 we characterize the bivariate distribution of self-confidence and ability by showing the frequency of each of the nine possible combinations of self-confidence and ability terciles in the representative NLSY79 sample. Recall that our model makes no assumption about this bivariate distribution, it simply predicts individuals' location decisions in their junior and senior periods conditional on self-confidence and ability. A first striking feature of the distribution is that individuals in the middle tercile of self-confidence are almost equally likely to be in the top (11.1%), middle (12.2%) and bottom (11.6%) terciles of ability. Not only are these percentiles similar, but they are also not far from the 11.1% that would correspond to a uniform bivariate distribution of self-confidence and ability. At the same time, it is relatively frequent for individuals to be in the top tercile of both self-confidence and ability (17.2% of the total) and relatively infrequent for individuals to be in the bottom tercile of self-confidence and in the top tercile of ability (5.1% of the total). Overall, the correlation between self-confidence and ability is 0.3.

In panel (b) of figure 2, using the same notation as in our theoretical figure 1, we assign to each grid cell the most prevalent location trajectory observed in the data for that combination of self-confidence and ability terciles. If individuals chose a location strategy independently of their ability and self-confidence, the prevalence of each location strategy in each of these nine cells should be the same regardless of ability and self-confidence. Instead, different location strategies turn out to be more or less prevalent depending on the values of ability and self-confidence.¹³

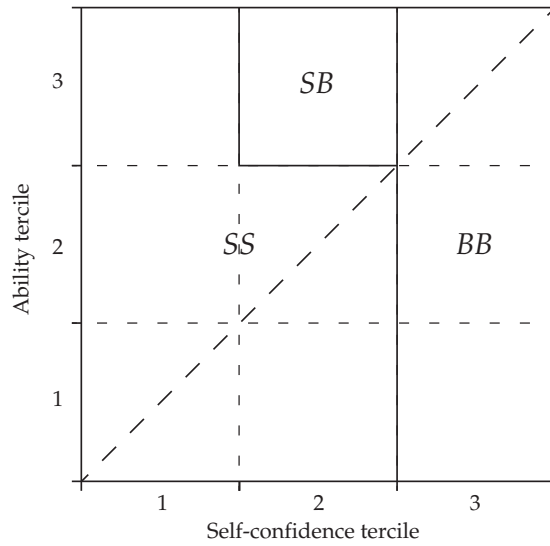
Looking first at the three cells along the diagonal of figure 2 (representing individuals whose self-confidence and ability are well aligned), we observe that individuals with low and intermediate values of both self-confidence and ability tend to locate in small cities when junior and to remain there. Similarly, individuals with high values of both self-confidence and ability tend to locate in big cities when junior and to remain there. This assortative matching between cities and workers with an accurate self-assessment matches well with our theoretical predictions. Looking back at the three cases in figure 1, we see that while all three have *SS* at the bottom-left corner and *BB* at the top-right corner, the complete coverage of the diagonal with *SS* and *BB* best matches case 2.

Turning to individuals whose self-assessment is less accurate, consider next individuals with intermediate values of self-confidence and high or low levels of ability. As already noted, such combinations of self-confidence and ability are very common in practice. Looking at panel (b) of figure 2, we notice that individuals with intermediate self-confidence and low ability (bottom-middle cell) tend to locate in small cities when junior and to remain there when senior. This matches the predictions of our model, where individuals with intermediate self-confidence choose to locate in a small city when junior and, subsequently, realizing their ability is in fact somewhat lower than expected, they have no reason to alter this initial location choice. Individuals who instead have

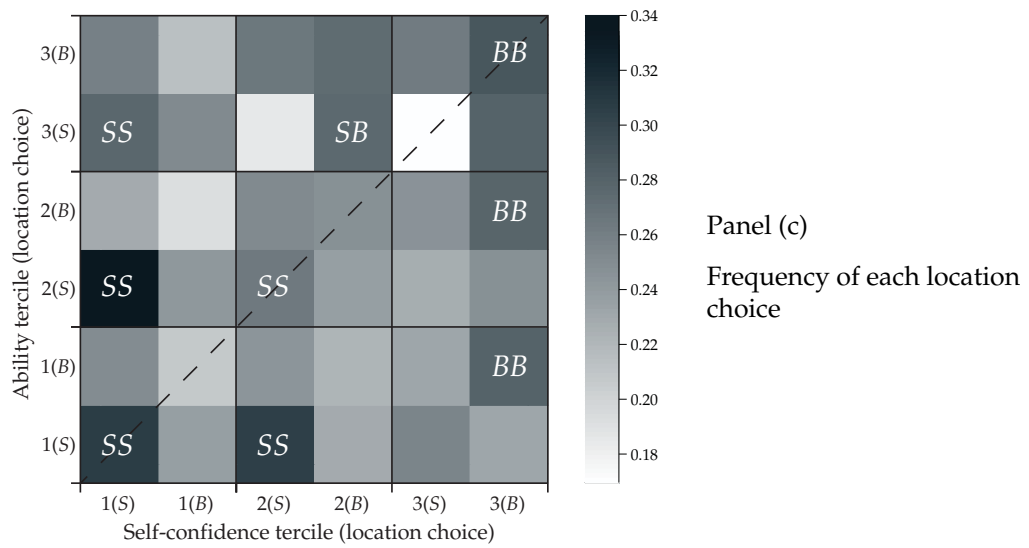
¹³As usual when measuring localisation, the relevant benchmark is not a uniform distribution but the distribution that would arise under random location choices (see e.g. Ellison and Glaeser, 1997, Duranton and Overman, 2005). Thus, we measure the prevalence of each location trajectory relative to a random-location benchmark in which each individual followed each location strategy with the same probability as the share of that strategy in the aggregate population regardless of ability and self-confidence.



Panel (a)
Bivariate distribution of self-confidence and ability



Panel (b)
Prevalent location choice



Panel (c)
Frequency of each location choice

Figure 2: Observed location choices by self-confidence and ability tertiles

intermediate self-confidence and high ability (top-middle cell) tend to locate in small cities when junior and then to relocate to big cities. In our model, under cases 1 and 2, once these individuals find they have high ability, they move away from small cities where their underconfidence led them to locate initially. Thus, empirical outcomes for individuals with intermediate self-confidence and high or low ability also match our theoretical predictions.

We have four remaining cells left to examine in panel (b) of figure 2. Individuals with intermediate levels of ability and high or low levels of self-confidence also follow the location choices predicted by our model. Those with high self-confidence (middle-right cell) start in big cities and remain there. According to our model, having made the investment of locating in a big city during the junior period led by their high self-confidence and having acquired valuable experience as a result, it is then worthwhile to remain in a big city in order to put that experience to use. Instead, those with the same intermediate ability but low self-confidence (middle-left cell) always locate in small cities. Once again, according to our model, initial location choices, by affecting the value of acquired experience, become self-perpetuating when errors in self-assessment are moderate.

The final two cells in panel (b) of figure 2 correspond to extreme mistakes in self-assessment. As seen in panel (a), these mistakes are less prevalent in practice than moderate mistakes. While all three cases that arise in our theoretical model have much in common, so far it is case 2 that best matches the empirical outcomes. We therefore also examine these last two cells in light of case 2. Individuals in the bottom-right cell have high self-confidence and low ability. In our model, these individuals are led by their high self-confidence to locate in big cities during their junior period. Whether they remain there or relocate to a small city depends on the magnitude of $\alpha_{BB>BS}$, as shown by the diagram for case 2 in figure 1. The empirical outcome in the data suggests that $\alpha_{BB>BS}$, the ability threshold beyond which individuals with high self-confidence find it optimal to remain in big cities, is quite low and, hence, this low threshold increases the prevalence of BB relative to BS in the bottom-right cell. Given that $\alpha_{BB>BS} \equiv \frac{\Delta\gamma}{e_B \pi_2 \Delta\Omega}$, then e_B and $\Delta\Omega$ should be large. Thus, workers with low ability who are driven to the big city when junior due to overconfidence and get lucky in solving the simple task are able to accumulate significant experience e_B . Despite their low ability, given their valuable big-city experience and the greater opportunities that big cities provide, $\Delta\Omega$, they choose to remain there. Our regression results below provide additional support for this interpretation.

Turning finally to individuals in the lowest tercile of self-confidence and the highest tercile of ability, they tend to locate in small cities both periods, as indicated by SS in the top-left cell of panel (b) of figure 2. Of all nine cells, this is the one where empirical outcomes provide a weaker match for our theoretical predictions. The other eight cells in panel (b) of figure 2 match with the theoretical predictions for case 2 in figure 1. Sticking to that case, we would expect a mixture of SS and SB in this range with the prevalence of one over the other determined by the magnitude of the ability threshold $\alpha_{SB>SS}$ with respect to $\alpha_{BB>SS}$. However, given that SB is the most prevalent location strategy in the top-middle cell (for individuals with intermediate self-confidence and high

ability), *SB* should also show up in the top-left cell for consistency.¹⁴

Panel (b) of figure 2 only shows the most prevalent location choice in each of the nine cells combining self-confidence and ability terciles. In panel (c) of figure 2 we provide a richer description of the data by showing, in addition to the prevalent location choice, the incidence of all choices in each cell. Each of the nine cells is now split into four quadrants corresponding to every possible two-period location strategy (*SS* in the bottom left, *SB* in the bottom right, *BS* in the top left, and *BB* in the top right) with darker shades representing a higher frequency of that strategy compared to the overall population. The prevalent location choice in each of the nine cells is marked in white over the corresponding quadrant of that cell.¹⁵ We can then see that while *BB* is the prevalent location strategy for the three cells on the right, corresponding to individuals in the highest tercile of self-confidence, the prevalence of *BB* becomes stronger the higher the ability tercile, as would be expected. Along the three top cells, corresponding to individuals in the highest tercile of ability, location strategy *BB* gradually loses ground to *SB* and then this to *SS* as we move leftwards towards lower terciles of self-confidence. Note also that, while *SS* is the dominant location strategy in the top-left cell, *SB* (the theoretical prediction in case 2) is not far. Strategy *SS* then becomes more clearly dominant for lower terciles of ability, again as would be expected.

Overall, we find that the location choices of individuals in their junior and senior periods vary with self-confidence and ability in a way that closely matches our theoretical predictions. However, this is based on raw data without taking into account other characteristics and experiences of individuals. We next turn to incorporating these.

Determinants of location in big and small cities

We now test key implications of our model by examining whether self-confidence and ability have an effect on the location decisions of individuals across cities of different sizes in their junior and senior periods, while controlling for other drivers of location and mobility. Specifically, we estimate logit models to look at the determinants of locating in a small or a big city either when junior or senior. In table 1 we report exponentiated coefficients (odds ratios), so coefficients above one indicate a positive effect and coefficients below one indicate a negative effect. Standard errors are clustered at the metropolitan area level and ***, **, and * indicate that a coefficient is significantly different from 1 (where 1 corresponds to an odds ratio implying no effect) at the 1, 5, and 10 percent levels.

A first implication of our model is that junior workers sort on self-confidence instead of on ability, so that we should expect more confident workers to have a higher probability of locating in big cities initially. In column (1) of table 1 we estimate a logit model where the dependent variable takes value one if the individual lives in a big city during his or her junior period, i.e. the year

¹⁴We get a complete match between theoretical predictions and empirical outcomes using alternative time frames for the senior period (such as an eight-year instead of a ten-year period). However, rather than pushing different margins to get a complete fit we prefer to note that this is the only cell where empirical predictions are less stable and sometimes depart from the theoretical prediction. This is not that surprising, given that this cell also has the fewest observations. Workers with self-confidence in the lowest tercile and ability in the highest tercile are relatively uncommon and tend to have a different type of low self-confidence: they tend to be simultaneously aware of their abilities and quite critical of themselves (Kohn and Schooler, 1969, Rosenberg, Schooler, Schoenbach, and Rosenberg, 1995).

¹⁵We are grateful to Jesse Shapiro, our discussant at the NBER Summer Institute, for suggesting this additional panel.

after completing education. Results show that individuals with higher levels of self-confidence are more likely to locate in a big city when junior. The corresponding coefficient on the self-confidence percentile (statistically significantly different from 1 at the 1% level) reveals that an increase of one standard deviation in the self-confidence percentile (28.6 points) raises the probability of locating in a big city by 12%.¹⁶ Instead, individuals with higher levels of ability are not any more likely to locate initially in a big city.

Our baseline model does not consider education, but our extension in appendix C does and predicts some sorting by educational attainment. Results reveal that having college education increases the probability of locating in a big city when junior by 112% (calculated by subtracting 1 from the estimated coefficient 2.122) relative to having at most primary education. We also include a set of conventional demographic controls. Having children reduces the probability of locating in a big city by 36%, and hispanics are also more likely to live in a big city during their junior period. Since many people are closely attached to the place where they grew up, we include as a control an indicator variable that takes value one if the individual was living in a small city at age 14. This turns out to be a crucial determinant of location decisions, greatly decreasing the probability of locating in a big city upon completing education and (as we shall see below) greatly increasing the probability of moving back home for those that do locate in a big city when junior. However, our results regarding the role of ability and self-confidence are very similar whether we control for location at age 14 or not.

Turning to the senior period, our model implies that ability—revealed after years of labour market experience—should matter more for the location of senior workers, although sorting on ability can still be quite imperfect. Thus, some successful high-ability workers should relocate from small to big cities while some unsuccessful low-ability workers should relocate from big to small cities. Further, conditional on the initial location choice, self-confidence should no longer drive location decisions in the senior period. To test these implications, we estimate two logit models in which the dependent variable captures a relocation across cities of different sizes between the junior and senior periods.

In column (2) of table 1 we focus exclusively on workers who were living in a small city upon completing education (hence the smaller number of observations relative to column 1) and study whether they moved subsequently. The dependent variable takes value one if the individual has relocated to a big city when observed in his or her senior period and value zero if the individual remains in a small city. Therefore, we examine the determinants of relocating from a small to a big city between periods. Results show that the level of self-confidence no longer influences the decision to relocate whereas the level of ability is a crucial relocation driver from small to big cities. The estimated coefficient implies that a one standard deviation increase in the ability percentile (28.2 points) raises the probability of moving to a big city by 21%. Thus, ability matters for the location of senior workers and, among the set of residents in small cities, it is the most able who move to a big city when senior. Other determinants show that having children or having lived in a small city at age 14 deter individuals located in a small city in their junior period from moving

¹⁶This figure is calculated by subtracting 1 from the estimated coefficient for the self-confidence percentile and multiplying this by the standard deviation of the variable: $(1.00433 - 1) \times 28.627 = 0.124$.

Table 1: Logit estimation of the determinants of location in big and small cities

	Probability of living in big city upon completing education	For individuals living in small city upon completing education, probability of having moved to big city 10 years later	For individuals living in big city upon completing education, probability of having moved to small city 10 years later
	(1)	(2)	(3)
Self-confidence percentile	1.004 (0.002)***	1.002 (0.002)	0.996 (0.003)
Cognitive ability percentile	1.000 (0.002)	1.007 (0.003)**	1.000 (0.004)
Male	0.909 (0.080)	0.922 (0.125)	1.155 (0.183)
Hispanic	2.003 (0.722)*	1.614 (0.820)	0.612 (0.240)
Black	1.395 (0.357)	1.076 (0.348)	0.490 (0.168)**
High-school graduate	0.915 (0.132)	0.827 (0.212)	0.867 (0.225)
Some college	0.955 (0.157)	1.413 (0.422)	0.818 (0.234)
College graduate	2.122 (0.519)***	1.857 (0.535)**	0.976 (0.306)
Never married	1.317 (0.227)	0.974 (0.241)	0.739 (0.167)
One or more children	0.642 (0.081)***	0.629 (0.103)***	1.133 (0.208)
Full-time working spouse	1.016 (0.161)	1.340 (0.149)***	0.828 (0.148)
Living in small city at age 14	0.018 (0.004)***	0.325 (0.078)***	2.408 (0.595)***
% time spent not working		1.005 (0.003)	1.011 (0.004)***
Relative wage		1.156 (0.148)	0.853 (0.151)
N	5,255	2,708	1,652
Pseudo R^2	0.463	0.074	0.055

Notes: Odd ratios (exponentiated coefficients) are reported, with coefficients above one indicating a positive effect and coefficients below one indicating a negative effect. All specifications include a constant and birth-year indicators. Standard errors in parentheses are clustered at the metropolitan area level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. A 'big city' is defined as a Core Based Statistical Area (CBSA) with a population greater than 2,000,000 in 2010. White, female, ever married and high-school dropouts are the omitted categories. Relative wage is the actual pre-migration wage of the individual relative to the wage predicted for the same individual by a regression including all the other variables in column (1), plus a city-size class fixed-effect, experience in each city-size class and its square, firm tenure and its square, three-digit occupation indicators, and three-digit sector indicators.

to a big city in their senior period. On the contrary, those individuals with college education or partnered with a full-time working spouse are more likely to make that move.

In column (3) of table 1 we focus exclusively on workers who were living in a big city during their junior period and study whether they moved subsequently.¹⁷ The dependent variable takes value one if the individual has relocated to a small city when observed in his or her senior period and value zero if the individual remains in a big city. Therefore, we examine the determinants of relocating from a big to a small city between periods. Results reveal that neither self-confidence nor ability are key determinants of the relocation decision of senior workers from big to small cities. This is consistent with one of our conclusions from figure 2: if big-city experience is highly valuable and differences in opportunities between cities are large, then workers who located in a big city when junior and were fortunate enough to have a positive experience there, tend to stay even if their ability is low.

Our model also suggests that workers who located in a big city when junior but had poor outcomes will subsequently relocate to a small city. These poor outcomes can be the result of luck or the worker being unable to benefit from the larger set of opportunities in big cities. We proxy these possibilities with two variables: the fraction of the time the individual spent not working in the big city since his or her junior period, and the residual of his or her wage that cannot be explained by observable wage determinants.¹⁸ Our findings show that the share of the time spent idle in a big city affects the likelihood of moving to a small city. The coefficient implies that a one standard deviation increase in the share of time spent not working (23.8%) increases the probability of moving to a small city by 27%. While the time spent idle has also a positive (though not statistically significant) effect on the probability of moving from a small to a big city, the magnitude of this effect is substantially smaller at 14%. The coefficient on the relative wage (i.e., the wage component that can not be explained by standard wage determinants) is not statistically significant, yet, it suggests that workers in the big city with earnings below their expected performance (i.e., with smaller wage residuals) are also more likely to move to a small city.¹⁹ In sum, our findings hint that unsuccessful workers in big cities tend to move to small cities, yet, they are not necessarily the least able workers among the pool of workers in big cities. In addition to relocations to small cities driven by poor performance in big cities, we also see that return migration is important: having lived in a small city at age 14 makes moving away from a

¹⁷The number of observations in columns (2) and (3) add up to the number of workers for whom we observe their senior period location and for whom we can construct all labor market controls, 4,360. This is lower than the number of observations in column (1) due to sample attrition over the ten years that separate both periods.

¹⁸For individuals who moved across cities of different sizes, we compute both variables at the last observation prior to the move, i.e., the last year the individual lived in a big city before moving to a small city when senior. For non-movers, we calculate both variables at a mid-term point between their junior and senior periods, i.e., we add five years to the junior period. To construct the fraction of the time spent not working, we add all periods the individual spent in unemployment or out of the labor force. To obtain wage residuals, we run a regression that includes all controls in column (1) of table 1, plus experience and its square, firm tenure and its square, a big city indicator variable, experience accumulated in big cities and its square, and occupation and sector indicators.

¹⁹On the contrary, in column (2) of table 1, point estimates imply that workers in small cities with earnings above their expected performance (i.e., higher wage residuals) are more likely to move to a big city.

big city more than twice as likely.²⁰

Robustness and alternative explanations

A key prediction of our framework is that young individuals with higher levels of self-confidence tend to be overrepresented in big cities. In our model, this is because individuals with high self-confidence see themselves as particularly able and ability is more highly rewarded in big cities. An alternative interpretation is that the sorting by self-confidence that we observe during the junior period is not related to an inaccurate assessment of ability, but instead reflects an intrinsic additional value of self-confidence in big cities. To address this concern, we now explicitly test whether self-confidence reinforces the advantages of big cities.

Our strategy builds on De la Roca and Puga (2017), who show that high ability is more highly rewarded in big than in small cities, as reflected in steeper earnings profiles. In an analogous way, we examine whether more self-confident workers also exhibit steeper wage profiles in big cities. We present results of this exercise in table 2. We begin in column (1) with a simplified version of the main estimation in De la Roca and Puga (2017) and regress log earnings on worker fixed-effects, time-varying job characteristics (measures of tenure, three-digit occupation and three-digit sector indicators), an indicator variable for currently working in a big city, measures of overall work experience and work experience acquired in big cities, and interactions between these measures of experience and worker fixed-effects.²¹

The positive and significant coefficient on big-city experience shows that experience acquired in big cities is significantly more valuable than experience acquired elsewhere, consistent with what we assume in our model. Specifically, comparing the coefficients on overall experience and experience acquired in big cities indicates that the first year of experience is 57% more valuable if accumulated in a big city instead of a small one.

Worker fixed-effects capture any time-invariant worker characteristics that are valuable in the job market. As in De la Roca and Puga (2017), an interaction between big-city experience and worker fixed-effects enters with a positive and significant coefficient in the regression, which they interpret as evidence that there is a positive interaction between ability and the long-term benefits of acquiring work experience in a big city. Since in our data we have a direct measure of cognitive ability in the AFQT percentile score, in column (2) we interact big-city experience with cognitive ability instead of the worker fixed effect. The positive and significant coefficient shows that cognitive ability, as measured by the AFQT percentile score, is more highly rewarded in big cities.

²⁰The large effect that having lived in a small city at age 14 has on the odds of returning (or moving to another small city) is in line with the findings in Kennan and Walker (2011) who emphasize the prevalence of return migration in the NLSY data. Return migration accounts for a large share of internal moves in the United States and Europe (DaVanzo, 1983, Greenwood, 1997, Hunt, 2004).

²¹We construct time-consistent three-digit occupation codes using crosswalks provided in Autor and Dorn (2013). We also construct time-consistent three-digit sector codes using the IPUMS consistent long-term classification 1990 (Ruggles, Genadek, Goeken, Grover, and Sobek, 2015). We consider only earnings observations for ages 23 and over since earlier employment spells are often just part-time jobs while studying.

Table 2: Estimation of the relationship of earnings, ability and self-confidence

	Dep. variable: Log earnings			
	(1)	(2)	(3)	(4)
Big city	.0762 (.0126)***	.0786 (.0127)***	.0765 (.0126)***	.0791 (.0127)***
Big-city experience	.0151 (.0027)***	.0157 (.0027)***	.0153 (.0027)***	.0157 (.0027)***
Big-city experience \times experience	-.0004 (.00009)***	-.0004 (.00009)***	-.0004 (.00009)***	-.0004 (.00009)***
Big-city experience \times worker fixed-effect	.0286 (.0156)*		.0309 (.0165)*	
Big-city exp \times experience \times worker fixed-effect	-.0005 (.0005)		-.0005 (.0006)	
Big-city experience \times cognitive ability		.0182 (.0100)*		.0206 (.0102)**
Big-city exp \times experience \times cognitive ability		-.0003 (.0003)		-.0003 (.0003)
Big-city experience \times self-confidence			-.0064 (.0102)	-.0070 (.0100)
Big-city exp \times experience \times self-confidence			.0001 (.0004)	.00007 (.0003)
Experience	.0265 (.0018)***	.0256 (.0018)***	.0265 (.0017)***	.0257 (.0018)***
Experience ²	-.0003 (.00006)***	-.0003 (.00006)***	-.0003 (.00006)***	-.0003 (.00006)***
Experience \times worker fixed-effect	.0723 (.0099)***		.0708 (.0106)***	
Experience ² \times worker fixed-effect	-.0013 (.0003)***		-.0013 (.0004)***	
Experience \times cognitive ability		.0393 (.0061)***		.0368 (.0063)***
Experience ² \times cognitive ability		-.0008 (.0002)***		-.0007 (.0002)***
Experience \times self-confidence			.0035 (.0063)	.0077 (.0061)
Experience ² \times self-confidence			-.00005 (.0002)	-.0001 (.0002)
Tenure	.0279 (.0013)***	.0282 (.0013)***	.0279 (.0013)***	.0282 (.0013)***
Tenure ²	-.0010 (.00006)***	-.0010 (.00006)***	-.0010 (.00006)***	-.0010 (.00006)***
3-digit occupation indicators & sector indicators	Yes	Yes	Yes	Yes
Worker fixed-effects	Yes	Yes	Yes	Yes
Observations	48,046	48,047	48,046	48,047
R ²	.2743	.2666	.2744	.2668

Notes: All specifications include a constant. Coefficients are reported with robust standard errors in parenthesis, which are clustered by worker. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. A 'big city' is defined as a Core Based Statistical Area (CBSA) with population greater than 2,000,000 in 2010. Worker fixed-effects computation follows De la Roca and Puga (2017).

In column (3), we add to the specification of column (1) interactions between both experience types and the Rosenberg self-confidence measure. The key coefficient of interest here is the one on the interaction between big-city experience and self-confidence. The fact that it is quantitatively small and not statistically significant indicates that, while earnings grow faster in big cities than in small cities, the effect is not stronger for more self-confident workers.²² This provides support for our interpretation that young workers sort according to their self-confidence not because self-confidence itself is more highly rewarded in big cities, but because high self-confidence reflects a high self-assessment of ability, and ability is more highly rewarded in big cities. Column (4) repeats the exercise adding interactions between both experience types and the Rosenberg self-confidence measure to the specification of column (2) instead of column (1). Compared with column (3), in column (4) we interact big-city experience with cognitive ability instead of the worker fixed effect, and reach similar conclusions.

Another concern is that our self-confidence measure may be capturing other relevant aspects of personality. We have seen that, as predicted by our model, individuals with high self-confidence are more likely to locate in a big city when young. However, high self-confidence may partly reflect other personality traits that could make a person more predisposed towards living in a big city. In particular, high self-confidence tends to be positively related to extraversion (Robins, Tracy, Trzesniewski, Potter, and Gosling, 2001). In turn, extravert individuals may be more likely to choose dense locations where they will tend to be less socially isolated. This pattern, where those who locate in urban areas as opposed to rural areas tend to be more extraverted, has been observed in some studies of the location choices of doctors and clergy (Francis and Rutledge, 2004, Jones *et al.*, 2013). Other studies, however, argue that the relationship between extraversion and location preferences is not as clear-cut, since big cities also favour anonymity which may help attract more introvert individuals. For instance, Marshall (1970) finds that measures of introversion at the individual level are highly correlated with a preference for privacy but not with a preference for solitude; and one of the strongest correlates of a preference for privacy is the size of the city or town an individual lives in, with more private individuals being significantly more likely to live in a bigger city. In any case, we would like to check that the relationship between self-confidence and the probability of living in a big city upon completing education is not driven by other personality traits correlated with self-confidence. We thus re-estimate our logistic regressions including measures of personality traits as additional controls.

Personality is most commonly assessed using a taxonomy of traits known as the big-five: extraversion, agreeableness, conscientiousness, neuroticism and openness to experiences. Unfortunately, the NLSY79 does not assess the big-five personality traits of its respondents. However, a related panel data set, the NLSY79 Children and Young Adults, does. This is a separate survey conducted to all offspring of NLSY79 female respondents. As part of this survey, young adults were administered the Ten Item Personality Inventory (TIPI) test, a ten-item questionnaire that measures the big five personality traits. In addition they were also subject to the same Rosenberg

²²We also examine whether more self-confident workers benefit from a level effect in earnings when moving to a big city by including an interaction between a big-city indicator variable and self-confidence. The coefficient on this interaction is not statistically significant.

Table 3: Logit estimation of the determinants of location in big cities with other personality traits

	Children of NLSY79	
	Probability of living in big city upon completing education	
	(1)	(2)
Self-confidence percentile	1.005 (0.002)**	1.005 (0.002)**
Math ability percentile	0.998 (0.003)	0.998 (0.003)
Reading recognition percentile	1.006 (0.003)*	1.006 (0.003)*
Reading comprehension percentile	0.997 (0.004)	0.997 (0.004)
Male	1.031 (0.124)	1.060 (0.119)
Hispanic	2.131 (0.744)**	2.152 (0.752)**
Black	1.613 (0.362)**	1.613 (0.372)**
High-school graduate	0.870 (0.154)	0.883 (0.154)
Some college	1.126 (0.255)	1.165 (0.256)
College graduate	3.443 (0.887)***	3.629 (0.933)***
Never married	1.059 (0.246)	1.057 (0.245)
One or more children	0.692 (0.112)**	0.698 (0.110)**
Working spouse	0.914 (0.186)	0.927 (0.186)
Living in small city at age 15	0.008 (0.002)***	0.008 (0.002)***
Extraversion percentile		0.999 (0.002)
Agreeableness percentile		1.002 (0.003)
Conscientiousness percentile		0.998 (0.002)
Emotional stability percentile		0.996 (0.003)
Openness to experiences percentile		1.003 (0.003)
Observations	3,161	3,161
Pseudo R^2	0.571	0.572

Notes: Odd ratios (exponentiated coefficients) are reported, with coefficients above one indicating a positive effect and coefficients below one indicating a negative effect. All specifications include a constant and birth-year indicators. Standard errors in parentheses are clustered at the metropolitan area level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. A 'big city' is defined as a Core Based Statistical Area (CBSA) with a population greater than 2,000,000 in 2010. White, female, ever married and high-school dropouts are the omitted categories. Math, reading recognition and reading comprehension percentiles use results from Peabody International Achievement tests. Personality percentiles are obtained using the Ten-Item Personality Inventory (TIPI) measure.

test that we use to measure self-confidence in our main sample. While NLSY79 Children and Young Adults respondents were not subject to the AFQT test that we use to measure cognitive ability in our main sample, they were administered various other cognitive ability tests, in particular the Peabody International Achievement Test (PIAT) for math, reading recognition and reading comprehension. Given that the AFQT combines four sections of the ASVAB test that measure math knowledge, arithmetic reasoning, paragraph comprehension and word knowledge, we include in our estimation the PIAT percentile scores for math, reading recognition and reading comprehension as measures of cognitive ability in place of the AFQT percentile score.²³

Since the NLSY79 Children and Young Adults follows offspring of women in our main data set, individuals are much younger, which prevents us from performing a full replication of our results with these data. In particular, the vast majority of respondents are too young to study the determinants of their location choices ten years after completing their education.²⁴ Hence, we cannot estimate the specifications for senior location choices in columns (2) and (3) of table 1. We can, however, estimate the specification for junior location choices in column (1) of table 1. Results are reported in table 3. In column (1) we reproduce the specification of column (1) of table 1, with the three ability measures replacing the unavailable AFQT percentile scores. As before, self-confidence is a significant determinant of the probability of locating in a big city upon completing education with a similar magnitude as in table 1. Of the three ability measures, only reading recognition (arguably a measure of education rather than of intrinsic ability) is a statistically significant determinant of living in a big city upon completing education. In column (2), we add to the specification the big-five personality traits. This shows that self-confidence remains a significant determinant of the probability of locating in a big city upon completing education, even after controlling for the big-five personality traits. There is a small positive correlation of 0.11 between the measure of extraversion and the measure of self-confidence. However, extraversion does not significantly affect the probability of locating in a big city. Neither do the other personality traits that are part of the big five. And the magnitude of the coefficient on self-confidence remains unchanged.

A final related concern is that self-confidence may be capturing attitude towards risk, with more self-confident individuals perhaps more willing to take risks. To the extent that locating in a big city may be seen as a risky investment, that could be driving in part the relationship between self-confidence and junior location in a big city. The NLSY79 includes a measure of attitude towards risk.²⁵ Respondents are asked to grade their willingness to take risks on a scale from 0 to 10. We use the individual's percentile in this scale (with a higher percentile associated with a lower willingness to take risks) as a measure of risk aversion. Table D.1 in appendix D replicates table 1 adding to each column the risk aversion percentile. In column (1) we notice that risk aversion

²³Respondents in the NLSY79 Children and Young Adults are administered the ability tests several times during childhood, starting at age 6. We use the latest test, for which the median age is 14.

²⁴The median age of NLSY79 Children and Young Adults respondents in our most recent survey year is 27, compared with a median age of 49 for NLSY79 respondents.

²⁵Unfortunately, this risk measure was collected only very recently, in 2010, whereas our measures of ability and self-confidence were collected at the beginning of the survey, when most respondents were teenagers. This is not ideal, since the measure may be affected by the consequences of earlier location choices, and also leads to a drop in the number of observations. For these reasons, we do not use this as our baseline specification.

does not significantly affect the probability of locating in a big city upon completing education. Moreover, controlling for risk aversion does not alter the effect of self-confidence estimated in column (1) of table 1, which is not surprising given that the measures of self-confidence and risk aversion are not significantly correlated in our sample. In columns (2) and (3), we observe that more risk averse individuals are less likely to relocate in either direction, as can be seen when studying moves from a small to a big city in column (2) or moves from a big to a small city in column (3). This finding is consistent with earlier research examining the positive association between willingness to take risks and migration (Jaeger, Dohmen, Falk, Huffman, Sunde, and Bonin, 2010, Bauernschuster, Falck, Heblich, Suedekum, and Lameli, 2014). And yet, controlling for risk aversion does not alter other coefficients.

7. Conclusions

Flawed self-assessment of own ability can help explain the limited sorting of workers across cities of different sizes, even though bigger cities provide higher-ability workers with disproportionately better learning experience and richer opportunities to exploit such experience. The reason is that workers whose self-confidence at an early stage of their career is not aligned with their ability may make location decisions they would not have made if they had known their actual ability to start with. By the time they learn enough about their actual ability, those early decisions have had a lasting impact, reducing their incentives to move and affecting their lifetime earnings. We have formalized this argument using an overlapping generations model with sorting across cities by workers who differ in self-confidence and ability, derived location and relocation patterns by self-confidence and ability from the model, and shown that they are empirically relevant using data for the United States. Besides helping explain limited urban sorting, these findings also confirm the power of personality traits as predictors and as causes of economic success, even after controlling for education, experience and cognition.

In particular, self-confidence is more important than actual ability for the location decisions of young workers. For older workers, ability plays a stronger role in determining location, but the lasting impact of their earlier choices limits the scope for relocation. Thus, some overconfident young workers start their career in a big city, while they would have chosen a small one had they correctly self-assessed their actual ability. If they nevertheless are fortunate enough to gain valuable experience, they tend to find that they can fully exploit this only by remaining in the big city also when older. Their initial misjudged decision thus becomes self-validating. Analogously, some underconfident young workers end up spending their whole life in a small city, even though a correct initial assessment of their ability would have made them self-select into a big city instead. Workers who seriously underestimate their own ability may nevertheless relocate from a small to a big city once their labour market experience provides them with better information of their true capabilities. Relocations from big to small cities appear to be driven instead by lack of success in the big city rather than by corrections to flawed self-assessment. Young workers who are confident enough of their own abilities locate in bigger cities to pursue their dreams, but those dreams do not come true for everyone.

References

- Almlund, Mathilde, Angela L. Duckworth, James J. Heckman, and Tim D. Kautz. 2011. Personality psychology and economics. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (eds.) *Handbook of the Economics of Education*, volume 4. Amsterdam: North-Holland, 1–181.
- Alonso, William. 1964. *Location and Land Use; Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.
- Autor, David and David Dorn. 2013. The growth of low skill service jobs and the polarization of the us labor market. *American Economic Review* 103(5): 1553–1597.
- Bacolod, Marigee, Bernardo Blum, and William C. Strange. 2009. Skills and the city. *Journal of Urban Economics* 65(2): 127–135.
- Bacolod, Marigee, Bernardo Blum, and William C. Strange. 2010. Elements of skill: Traits, intelligences, education, and agglomeration. *Journal of Regional Science* 50(1): 245–280.
- Bauernschuster, Stefan, Oliver Falck, Stephan Heblich, Jens Suedekum, and Alfred Lameli. 2014. Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior & Organization* 98: 56–69.
- Baum-Snow, Nathaniel and Ronni Pavan. 2012. Understanding the city size wage gap. *Review of Economic Studies* 79(1): 88–127.
- Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy* 122(3): 507–553.
- Benabou, Roland and Jean Tirole. 2002. Self-confidence and personal motivation. *Quarterly Journal of Economics* 117(3): 871–915.
- Berry, Steven and Edward L. Glaeser. 2005. The divergence in human capital levels across cities. *Papers in Regional Science* 84(3): 407–444.
- Brueckner, Jan K. 1987. The structure of urban equilibria: A unified treatment of the Muth-Mills model. In Edwin S. Mills (ed.) *Handbook of Regional and Urban Economics*, volume 2. Amsterdam: North-Holland, 821–845.
- Chen, Gilad, Stanley M. Gully, and Dov Eden. 2001. Validation of a new general self-efficacy scale. *Organizational Research Methods* 4(1): 62–83.
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2016. The costs of agglomeration: House and land prices in French cities. Processed, Wharton School, University of Pennsylvania.
- DaVanzo, Julie. 1983. Repeat migration in the United States: Who moves back and who moves on? *Review of Economics and Statistics* 65(4): 552–59.
- Davis, Donald R. and Jonathan I. Dingel. 2012. A spatial knowledge economy. Working Paper 18188, National Bureau of Economic Research.
- Davis, Donald R. and Jonathan I. Dingel. 2013. The comparative advantage of cities. Processed, Columbia University.
- De la Roca, Jorge and Diego Puga. 2017. Learning by working in big cities. *Review of Economic Studies* 84(1): 106–142.

- Dunning, David, Chip Heath, and Jerry M. Suls. 2004. Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science In The Public Interest* 5(3): 69–106.
- Duranton, Gilles and Henry G. Overman. 2005. Testing for localization using micro-geographic data. *Review of Economic Studies* 72(4): 1077–1106.
- Duranton, Gilles and Diego Puga. 2001. Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review* 91(5): 1454–1477.
- Duranton, Gilles and Diego Puga. 2015. Urban land use. In Gilles Duranton, Vernon Henderson, and William Strange (eds.) *Handbook of Regional and Urban Economics*, volume 5. Amsterdam: North-Holland, 467–560.
- Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. Spatial sorting. *Journal of Political Economy* 122(3): 554–620.
- Ellison, Glenn and Edward L. Glaeser. 1997. Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy* 105(5): 889–927.
- Francis, Leslie J. and Christopher J. F. Rutledge. 2004. Personality and preference for rural ministry: Replication and reconsideration. *Pastoral Psychology* 53(1): 43–48.
- Glaeser, Edward L. 1999. Learning in cities. *Journal of Urban Economics* 46(2): 254–277.
- Glaeser, Edward L. and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19(2): 316–342.
- Greenwood, Michael J. 1997. Internal migration in developed countries. In Mark R. Rosenberg and Oded Stark (eds.) *Handbook of Population and Family Economics*, volume 1B. Amsterdam: North-Holland, 647–720.
- Hansford, Brian C. and John A. Hattie. 1982. The relationship between self and achievement/performance measures. *Review of Educational Research* 52(1): 123–142.
- Harter, Susan. 1990. Causes, correlates and the functional role of global self-worth: A life-span perspective. In John Kolligian and Robert J. Sternberg (eds.) *Competence Considered*. New Haven, CT: Yale University Press, 67–98.
- Hunt, Jennifer. 2004. Are migrants more skilled than non-migrants? Repeat, return, and same-employer migrants. *Canadian Journal of Economics* 37(4): 830–849.
- Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2010. Direct evidence on risk attitudes and migration. *Review of Economics and Statistics* 92(3): 684–689.
- Jones, Michael P., Diann Eley, Lisa Lampe, Carissa M. Coulston, Gin S. Malhli, Ian Wilson, Brian Kelly, Nicholas J. Talley, Cathy Owen, Gerry Corrigan, Barbara Griffin, John Humphreys, Beatrice Alba, and Pamela Stagg. 2013. Role of personality in medical students' initial intention to become rural doctors. *Australian Journal of Rural Health* 21(2): 80–89.
- Judge, Timothy A., Amir Erez, and Joyce E. Bono. 1998. The power of being positive: The relation between positive self-concept and job performance. *Human Performance* 11(2/3): 167–187.
- Kennan, John and James R. Walker. 2011. The effect of expected income on individual migration decisions. *Econometrica* 79(1): 211–251.

- Kohn, Melvin L. and Carmi Schooler. 1969. Class, occupation, and orientation. *American Sociological Review* 34(5): 659–678.
- Marshall, Nancy J. 1970. Personality correlates of orientation toward privacy. In *EDRA 2: Proceedings of the 2nd Annual Environmental Design Research Association Conference*. 316–319.
- Mills, Edwin S. 1967. An aggregative model of resource allocation in a metropolitan area. *American Economic Review Papers and Proceedings* 57(2): 197–210.
- Moretti, Enrico. 2012. *The New Geography of Jobs*. New York, NY: Houghton Mifflin Harcourt.
- Muth, Richard F. 1969. *Cities and Housing*. Chicago: University of Chicago Press.
- Robins, Richard W., Jessica L. Tracy, Kali Trzesniewski, Jeff Potter, and Samuel D. Gosling. 2001. Personality correlates of self-esteem. *Journal of Research in Personality* 35(4): 463–482.
- Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton: Princeton University Press.
- Rosenberg, Morris, Carmi Schooler, Carrie Schoenbach, and Florence Rosenberg. 1995. Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review* 60(1): 141–156.
- Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0*. Minneapolis, MN: University of Minnesota.
- Stajkovic, Alexander D. and Fred Luthans. 1998. Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin* 124(2): 240–261.
- Yankow, Jeffrey J. 2006. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics* 60(2): 139–161.

Appendix A. Proof of Proposition 1

First, note that the three cases are mutually exclusive. Since $e_S < e_B$ and $\Omega_S < \Omega_B$ it follows that $\frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma} < \frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma}$. Hence the condition for case 1, $\frac{\Delta e}{\Delta \Omega^2} < \frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma}$, and the condition for case 3, $\frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma} < \frac{\Delta e}{\Delta \Omega^2}$, cannot hold simultaneously. The condition for case 2, $\frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma} \leq \frac{\Delta e}{\Delta \Omega^2} \leq \frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma}$, is precisely that neither the condition for case 1 nor the condition for case 3 hold.

Case 1 arises when $\alpha_{SB \succ SS} < \alpha_{BB \succ SB}$. Substituting equations (7) and (15) into this inequality and rearranging leads to the condition $\frac{\Delta e}{\Delta \Omega^2} < \frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma}$.

Starting from the lowest ability, consider first workers with $\alpha \leq \alpha_{BB \succ BS}$. By Lemma 2, such workers always locate in S when senior, while they also locates in S when junior if and only if $\alpha \leq \alpha_{BS \succ SS}$. Since $e_S < e_B$, it is always true that $\alpha_{BB \succ BS} < \alpha_{SB \succ SS}$ so that $\alpha \leq \alpha_{BB \succ BS}$ implies $\alpha < \alpha_{SB \succ SS}$. The parameter condition defining case 1 is $\alpha_{SB \succ SS} < \alpha_{BB \succ SB}$. And since $\Omega_S < \Omega_B$, it is always true that $\alpha_{BB \succ SB} < \alpha_{BS \succ SS}$. Thus, in case 1, $\alpha \leq \alpha_{BB \succ BS}$ also implies $\alpha < \alpha_{BS \succ SS}$ so that workers with $\alpha \leq \alpha_{BB \succ BS}$ locate in S both periods.

Consider next workers with $\alpha_{BB \succ BS} < \alpha \leq \alpha_{SB \succ SS}$. The condition $\alpha \leq \alpha_{SB \succ SS}$ is equivalent to $U_{SB} - U_{SS} \leq 0$. The parameter condition defining case 1 is $\alpha_{SB \succ SS} < \alpha_{BB \succ SB}$, so that $\alpha \leq \alpha_{SB \succ SS}$ also implies $\alpha < \alpha_{BB \succ SB}$, equivalent to $U_{BB} - U_{SB} < 0$. Adding these two inequalities on utility

levels yields $U_{BB} - U_{SS} < 0$, which is equivalent to $\alpha < \alpha_{BB>SS}$. By Lemma 2, workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in S both periods if and only if $\alpha \leq \alpha_{BB>SS}$. Since in case 1 both workers with $\alpha \leq \alpha_{BB>BS}$ and workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in S both periods, we can pool them together stating that workers with $\alpha \leq \alpha_{SB>SS}$ locate in S both periods.

Moving up to intermediate ability, consider workers with $\alpha_{SB>SS} < \alpha \leq \alpha_{BB>SB}$. By Lemma 2, workers with $\alpha_{SB>SS} < \alpha$ locate in B when senior unless they fail at the simple task when junior, while they locate in S when junior if and only if $\alpha \leq \alpha_{BB>SB}$. Thus, workers with $\alpha_{SB>SS} < \alpha \leq \alpha_{BB>SB}$ locate in S when junior and, if and only if successful, relocate to B when senior.

To conclude case 1, consider workers with $\alpha_{BB>SB} < \alpha$. The parameter condition defining case 1 is $\alpha_{SB>SS} < \alpha_{BB>SB}$, so that $\alpha_{BB>SB} < \alpha$ also implies $\alpha_{SB>SS} < \alpha$. By Lemma 2, workers with $\alpha_{SB>SS} < \alpha$ locate in B when senior unless they fail at the simple task when junior, while they also locate in B when junior if and only if $\alpha_{BB>SB} < \alpha$. Hence, workers with $\alpha_{BB>SB} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Case 2 arises when $\alpha_{BB>SB} \leq \alpha_{SB>SS}$ and $\alpha_{BB>BS} \leq \alpha_{BS>SS}$ simultaneously. Substituting equations (3), (7), (10) and (15) into this inequality and rearranging leads to the condition $\frac{\pi_2 e_S^2}{\Omega_B \Delta \gamma} \leq \frac{\Delta e}{\Delta \Omega^2} \leq \frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma}$.

To examine case 2, it is useful to begin by showing that under the parameter conditions defining this case we must have $\alpha_{BB>BS} \leq \alpha_{BB>SS} \leq \alpha_{SB>SS}$. Suppose that, on the contrary, $\alpha_{SB>SS} < \alpha_{BB>SS}$. Then at $\alpha = \alpha_{SB>SS}$ we have $U_{SB} - U_{SS} = 0$ and $U_{BB} - U_{SS} < 0$. Subtracting these two inequalities on utility levels yields $U_{BB} - U_{SB} < 0$, which is equivalent to $\alpha < \alpha_{BB>SB}$. This would imply $\alpha_{SB>SS} < \alpha_{BB>SB}$ which contradicts the condition defining case 2 that $\alpha_{BB>SB} \leq \alpha_{SB>SS}$. Similarly, suppose that $\alpha_{BB>SS} \leq \alpha_{BB>BS}$. Then at $\alpha = \alpha_{BB>BS}$ we have $U_{BB} - U_{BS} = 0$ and $U_{BB} - U_{SS} > 0$. Subtracting these two inequalities on utility levels yields $U_{BS} - U_{SS} > 0$, which is equivalent to $\alpha_{BS>SS} < \alpha$. This would imply $\alpha_{BS>SS} < \alpha_{BB>BS}$ which contradicts the condition defining case 2 that $\alpha_{BB>BS} \leq \alpha_{BS>SS}$. Consequently, in case 2 $\alpha_{BB>BS} \leq \alpha_{BB>SS} \leq \alpha_{SB>SS}$. With this in mind, we now examine location choices in case 2 depending on α .

Starting from the lowest ability, consider first workers with $\alpha \leq \alpha_{BB>BS}$. By Lemma 2, such workers always locate in S when senior, while they also locates in S when junior if and only if $\alpha \leq \alpha_{BS>SS}$. One of the parameter conditions for case 2 is $\alpha_{BB>BS} \leq \alpha_{BS>SS}$, so that $\alpha \leq \alpha_{BB>BS}$ also implies $\alpha \leq \alpha_{BS>SS}$. Thus, workers with $\alpha \leq \alpha_{BB>BS}$ locate in S both periods.

Moving up to intermediate ability, consider workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$. By Lemma 2, if $\alpha \leq \alpha_{BB>SS}$ such workers locate in S both periods, while if $\alpha_{BB>SS} < \alpha$, they locate in B both periods unless they fail at the simple task when junior; if they do fail, then they relocate to S when senior.

To conclude case 2, consider workers with $\alpha_{SB>SS} < \alpha$. By Lemma 2, such workers locate in B when senior unless they fail at the simple task when junior, while they also locates in B when junior if and only if $\alpha_{BB>SB} < \alpha$. One of the parameter conditions for case 2 is $\alpha_{BB>SB} \leq \alpha_{SB>SS}$, so that $\alpha_{SB>SS} < \alpha$ also implies $\alpha_{BB>SB} < \alpha$. Thus, workers with $\alpha_{SB>SS} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Since in case 2 $\alpha_{BB>BS} \leq \alpha_{BB>SS} \leq \alpha_{SB>SS}$, location choices can be summarized by saying that

workers with $\alpha \leq \alpha_{BB>SS}$ locate in S both periods; while workers with $\alpha_{BB>SS} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

Case 3 arises when $\alpha_{BS>SS} < \alpha_{BB>BS}$. Substituting equations (3) and (10) into this inequality and rearranging leads to the condition $\frac{\pi_2 e_B^2}{\Omega_S \Delta \gamma} < \frac{\Delta e}{\Delta \Omega^2}$.

Starting from the lowest ability, consider first workers with $\alpha \leq \alpha_{BS>SS}$. The parameter condition defining case 3 is $\alpha_{BS>SS} < \alpha_{BB>BS}$, so that $\alpha \leq \alpha_{BS>SS}$ also implies $\alpha < \alpha_{BB>BS}$. By Lemma 2, workers with $\alpha \leq \alpha_{BB>BS}$ locate in S when senior, while they also locate in S when junior if and only if $\alpha < \alpha_{BS>SS}$. Thus, in case 3, workers with $\alpha \leq \alpha_{BS>SS}$ locate in S both periods.

Moving up to intermediate ability, consider workers with $\alpha_{BS>SS} < \alpha \leq \alpha_{BB>BS}$. By Lemma 2, workers with $\alpha \leq \alpha_{BB>BS}$ locate in S when senior, while they locate in B when junior if and only if $\alpha_{BS>SS} < \alpha$. Thus, workers with $\alpha_{BS>SS} < \alpha \leq \alpha_{BB>BS}$ locate in B when junior and relocate to S when senior.

Consider next workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$. The condition $\alpha_{BB>BS} < \alpha$ is equivalent to $U_{BB} - U_{BS} > 0$. The parameter condition defining case 3 is $\alpha_{BS>SS} < \alpha_{BB>BS}$, so that $\alpha_{BB>BS} < \alpha$ also implies $\alpha_{BS>SS} < \alpha$, equivalent to $U_{BS} - U_{SS} > 0$. Adding these two inequalities on utility levels yields $U_{BB} - U_{SS} > 0$, which is equivalent to $\alpha_{BB>SS} < \alpha$. By Lemma 2, workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locate in B both periods if $\alpha_{BB>SS} < \alpha$, unless they fail at the simple task when junior, in which case they relocate to S when senior.

To conclude case 3, consider workers with $\alpha_{SB>SS} < \alpha$. By Lemma 2, such workers locate in B when senior unless they fail at the simple task when junior, in which case they locate in S when senior. Also by Lemma 2, such workers locate in B when junior if and only if $\alpha_{BB>SB} < \alpha$. Since $e_S < e_B$, it is always the case that $\alpha_{BB>BS} < \alpha_{SB>SS}$ so that $\alpha_{SB>SS} < \alpha$ implies $\alpha_{BB>BS} < \alpha$. The parameter condition defining case 3 is $\alpha_{BS>SS} < \alpha_{BB>BS}$. Since $\Omega_S < \Omega_B$, it is always the case that $\alpha_{BB>SB} < \alpha_{BS>SS}$. Thus, in case 3, $\alpha_{SB>SS} < \alpha$ also implies $\alpha_{BB>SB} < \alpha$, so that workers with $\alpha_{SB>SS} < \alpha$ locate in B both periods if $\alpha_{BB>SS} < \alpha$, unless they fail at the simple task when junior, in which case they relocate to S when senior. Since in case 3 both workers with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ and workers with $\alpha_{SB>SS} < \alpha$ choose the same locations, we can pool them together stating that workers with $\alpha_{BB>BS} < \alpha$ locate in B both periods unless they fail at the simple task when junior, in which case they relocate to S when senior.

□

Appendix B. Proof of Proposition 3

Define the auxiliary function

$$\tilde{b}(n) = 1 + \frac{n}{2} - b(n). \quad (\text{B.1})$$

This is the difference between the population of B , $1 + \frac{n}{2}$, and the number of workers who wish to locate in B given that population, $b(n)$. Existence and uniqueness of the urban equilibrium can be proven by showing that $\tilde{b}(n)$ has a single root in the feasible interval $0 \leq n \leq 2$.

We begin by showing that $b(n)$ is a continuous decreasing function of n over the interval $[0,2]$. First, $b(n)$ is a continuous and decreasing function of n in each of the three open or half-open

intervals $[0, \underline{n}]$, (\underline{n}, \bar{n}) , and $(\bar{n}, 2]$. Consider $b(n)$ for $n \in [0, \underline{n}]$. In this interval $b(n)$ is continuous in n : by the fundamental theorem of calculus, it is a continuous function of $\alpha_{BB>SB}(n)$, $\alpha_{BB>BS}(n)$, and $\alpha_{SB>SS}(n)$, which are in turn continuous functions of n . From equation (22), by the fundamental theorem of calculus and the chain rule of derivation, its derivative with respect to n can be written

$$\begin{aligned}
b'(n) \Big|_{0 \leq n < \underline{n}} &= -\alpha'_{BB>SB}(n) \int_0^1 f(\alpha_{BB>SB}(n), \alpha) d\alpha - \alpha'_{BB>SB}(n) \int_{\alpha_{BB>BS}(n)}^{\alpha_{SB>SS}(n)} \alpha f(\alpha_{BB>SB}(n), \alpha) d\alpha \\
&\quad - \alpha'_{BB>BS}(n) \alpha_{BB>BS}(n) \int_{\alpha_{BB>SB}(n)}^1 f(\sigma, \alpha_{BB>BS}(n)) d\sigma \\
&\quad - \alpha'_{SB>SS}(n) \alpha_{SB>SS}(n) \int_0^{\alpha_{BB>SB}(n)} f(\sigma, \alpha_{SB>SS}(n)) d\sigma,
\end{aligned} \tag{B.2}$$

which is negative given that $\alpha'_{BB>SB}(n) > 0$, $\alpha'_{BB>BS}(n) > 0$ and $\alpha'_{SB>SS}(n) > 0$. The continuity of $b(n)$ over the intervals (\underline{n}, \bar{n}) and $(\bar{n}, 2]$ can be proven analogously.

Second, $b(n)$ is continuous in n also at $n = \underline{n}$ and $n = \bar{n}$. Consider the continuity of $b(n)$ at $n = \underline{n}$. In equation (22) we see that the only difference between $b(n)|_{0 \leq n < \underline{n}}$ and $b(n)|_{\underline{n} \leq n \leq \bar{n}}$ is that wherever $\alpha_{BB>SB}(n)$ appears in $b(n)|_{0 \leq n < \underline{n}}$, $\alpha_{BB>SS}(n)$ appears instead in $b(n)|_{\underline{n} \leq n \leq \bar{n}}$. Since $\alpha_{BB>SS}(\underline{n}) = \alpha_{SB>SS}(\underline{n})$ and since, by definition of \underline{n} , $\alpha_{SB>SS}(\underline{n}) = \alpha_{BB>SB}(\underline{n})$, it follows that

$$\lim_{n \rightarrow \underline{n}^-} b(n) = b(\underline{n}) = \lim_{n \rightarrow \underline{n}^+} b(n) \tag{B.3}$$

The continuity of $b(n)$ at $n = \bar{n}$ can be proven analogously.

Since $1 + \frac{n}{2}$ is a continuous increasing function of n and $b(n)$ is a continuous decreasing function of n over the interval $[0, 2]$, it follows that $\tilde{b}(n) = 1 + \frac{n}{2} - b(n)$ is a continuous increasing function of n over this interval.

By equation (20), $n = 0$ implies $\Delta\gamma = 0$; which in turn, by equations (3), (7), (10), (13), and (15), implies $\alpha_{BB>BS} = \alpha_{SB>SS} = \alpha_{BS>SS} = \alpha_{BB>SS} = \alpha_{BB>SB} = 0$; and substituting these into equation (22) yields $b(0) = 2$; which, by equation (B.1), implies $\tilde{b}(0) = -1$. Moreover, since $1 + \frac{n}{2}$ takes value 2 for $n = 2$, and since $b(n)$ is decreasing in n over the interval $[0, 2]$ starting from the value $b(0) = 2$, it follows that $\tilde{b}(2) > 0$

Since $\tilde{b}(n)$ is a continuous function of n over the interval $[0, 2]$, $\tilde{b}(0) < 0$, and $\tilde{b}(2) > 0$, by Bolzano's Theorem there exists at least one value of $n \in (0, 2)$ such that $\tilde{b}(n) = 0$. This proves that an urban equilibrium exists. In addition, both the big and small cities are populated in equilibrium (i.e., the equilibrium value of n satisfies $0 < n < 2$ with strict inequality). The urban equilibrium is also unique. Suppose on the contrary that there were two or more values of n in $(0, 2)$ such that $\tilde{b}(n) = 0$. Then, by Rolle's Theorem there would have to be some n in this interval such that $\tilde{b}'(n) = 0$, which contradicts our previous result that $\tilde{b}'(n) > 0$ over the interval $[0, 2]$.

Turning to comparative statics, totally differentiating the equilibrium condition $\tilde{b}(n) = 1 + \frac{n}{2} - b(n) = 0$ and solving for $\frac{dn}{d\tau}$ yields

$$\frac{dn}{d\tau} = \frac{\frac{db(n)}{d\tau}}{\tilde{b}'(n)}. \tag{B.4}$$

Since τ and n always enter $b(n)$ together as a product (because $\Delta\gamma$ enters every threshold level of α and, by equation 20, $\Delta\gamma = \tau n$), it follows that $\frac{db(n)}{d\tau} = b'(n)$, and we have already shown that $b'(n) < 0$. We have also shown that $\tilde{b}'(n) > 0$. Hence, we can sign equation (B.4): $\frac{dn}{d\tau} < 0$. The comparative statics $\frac{dn}{d\Delta\Omega} > 0$ and $\frac{dn}{d\Delta e} > 0$ can be proven analogously. \square

Appendix C. Introducing education and uncertainty over self-assessment

For simplicity, in the main text we do not consider workers' education. Also, we treat the situation in which self-confidence accurately reflects ability as one where each worker knows her ability α with certainty; and we treat the situation in which self-confidence does not reflect ability accurately as one where each junior worker is certain that her ability is σ (self-confidence) until she updates this belief to the actual value α while attempting to complete a simple task.

In this appendix, we introduce differences in education across workers. In particular, workers start their junior period with one of two education levels: high or low. We let this level of educational attainment be informative about workers' ability, although we do not explicitly model the process that results in different educational achievements. In addition, in this alternative formulation, we also let junior workers realize their self-assessment may be inaccurate.

As in the baseline model in the main text, we consider two cases: accurate and imperfect self-assessment. However, accurate self-assessment no longer implies that each junior worker knows her ability α with certainty. Instead, she receives a signal that allows her to compute an unbiased estimate of her ability, so she takes her ability to have expected value α and variance θ . The situation in which self-confidence does not reflect ability accurately is modelled in this alternative formulation by having each worker receive a signal that leads her to a possibly biased estimate of her ability. However, she realizes this self-assessment may be inaccurate, and she takes her ability to have expected value $\sigma \neq \alpha$ and variance θ . Education is informative about a worker's ability, so that σ takes on average a higher value for workers with higher education. This formulation finds empirical support in our data. According to our self-confidence measure, workers with a college degree score in the 64th percentile on average compared to a score in the 50th percentile for workers with lower education. Further, the variance in scores does not differ by educational attainment. We still maintain the assumption that while trying to complete their simple task workers learn about their true ability, so all senior workers know their α accurately.

Since senior workers know their ability, the senior location choice remains exactly as in the main text and follows lemma 1, with the two thresholds on which senior-period location depends, $\alpha_{BB>BS}$ and $\alpha_{SB>SS}$, unchanged.

For the junior location choice under accurate self-assessment, we need to consider the same three ranges of ability as before: $\alpha \leq \alpha_{BB>BS}$, $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$, and $\alpha_{SB>SS} < \alpha$. For a worker with low ability $\alpha \leq \alpha_{BB>BS}$ locating in B when senior is never worthwhile regardless of her junior period location. She locates in B in her junior period if and only if

$$E[U_{BS}(\alpha) - U_{SS}(\alpha)] = (\alpha^2 + \theta)\Omega_S(e_B - e_S)\pi_2 - (\gamma_B - \gamma_S) > 0, \quad (C.1)$$

or, equivalently, if and only if

$$\alpha > \alpha_{BS>SS} \equiv \sqrt{\frac{\Delta\gamma}{\Omega_S \pi_2 \Delta e}} - \theta. \quad (\text{C.2})$$

Compared with the value of $\alpha_{BS>SS}$ in equation (10), the only difference is the term $-\theta$. Uncertainty over ability and convexity lower the (expected) ability threshold above which a junior worker locates in B .

For a worker with $\alpha_{BB>BS} < \alpha \leq \alpha_{SB>SS}$ locating in B when senior is worthwhile if she located in B when junior and successfully completed her simple task. Knowing this, she locates in B in her junior period if and only if

$$E [U_{BB}(\alpha) - U_{SS}(\alpha)] = (\alpha^2 + \theta)(\Omega_B e_B - \Omega_S e_S) \pi_2 - (1 + \alpha)(\gamma_B - \gamma_S) > 0, \quad (\text{C.3})$$

or, equivalently, if and only if

$$\alpha > \alpha_{BB>SS} \equiv \frac{1}{2} \left(\tilde{\alpha} + \sqrt{\tilde{\alpha}^2 + 4(\tilde{\alpha} - \theta)} \right), \quad (\text{C.4})$$

where $\tilde{\alpha}$ is still defined as in equation (12). Compared with the value of $\alpha_{BB>SS}$ in equation (13), the only difference is the term $-\theta$.

For a worker with high ability $\alpha_{SB>SS} < \alpha$ locating in B when senior (conditional on junior period success) is always worthwhile regardless of her junior period location. Knowing she will locate in B when senior if successful, she locates in B in her junior period if and only if

$$E [U_{BB}(\alpha) - U_{SB}(\alpha)] = (\alpha^2 + \theta)\Omega_B(e_B - e_S) \pi_2 - (\gamma_B - \gamma_S) > 0, \quad (\text{C.5})$$

or, equivalently, if and only if

$$\alpha > \alpha_{BB>SB} \equiv \sqrt{\frac{\Delta\gamma}{\Omega_B \pi_2 \Delta e}} - \theta. \quad (\text{C.6})$$

Compared with the value of $\alpha_{BB>SB}$ in equation (14), the only difference is the term $-\theta$.

Lemma 2 still applies as before, but with the thresholds $\alpha_{BS>SS}$, $\alpha_{BB>SS}$, and $\alpha_{BB>SB}$ now defined by equations (C.2), (C.4), and (C.6) respectively instead of by equations (10), (13), and (15).

The three cases in propositions 1 and 2 need to be redefined using the new values of the thresholds. Case 1 arises when $\alpha_{SB>SS} < \alpha_{BB>SB}$. Substituting equations (7) and (C.6) into this inequality and rearranging leads to the updated parameter condition defining case 1: $\frac{\Delta e}{\Delta\Omega^2} < \frac{\pi_2 e_S^2}{\Omega_B \Delta\gamma} - \frac{\theta e_S^2 \pi_2^2 \Delta e}{\Delta\gamma^2}$. Case 2 arises when $\alpha_{BB>SB} \leq \alpha_{SB>SS}$ and $\alpha_{BB>BS} \leq \alpha_{BS>SS}$ simultaneously. Substituting equations (3), (7), (C.2) and (C.6) into this inequality and rearranging leads to the updated parameter condition defining case 2: $\frac{\pi_2 e_S^2}{\Omega_B \Delta\gamma} - \frac{\theta e_S^2 \pi_2^2 \Delta e}{\Delta\gamma^2} \leq \frac{\Delta e}{\Delta\Omega^2} \leq \frac{\pi_2 e_B^2}{\Omega_S \Delta\gamma} - \frac{\theta e_B^2 \pi_2^2 \Delta e}{\Delta\gamma^2}$. Case 3 arises when $\alpha_{BS>SS} < \alpha_{BB>BS}$. Substituting equations (3) and (C.2) into this inequality and rearranging leads to the updated parameter condition defining case 3: $\frac{\pi_2 e_B^2}{\Omega_S \Delta\gamma} - \frac{\theta e_B^2 \pi_2^2 \Delta e}{\Delta\gamma^2} < \frac{\Delta e}{\Delta\Omega^2}$.

The proof of proposition 1 in appendix A in the main text is still valid with the updated values for the thresholds $\alpha_{BS>SS}$, $\alpha_{BB>SS}$, and $\alpha_{BB>SB}$ and the updated parameter conditions for each of the three cases. The same is true about the proof of proposition 2, for the case with imperfect

self-assessment. Thus all of our results are qualitatively unchanged. There are two differences introduced by this extension of the model.

The first difference is that the (perceived) ability thresholds above which junior workers locate in B are lower with uncertainty since utility is convex in ability (which matters both periods, creating a positive quadratic term in α). Furthermore, we can see from the updated parameter conditions for each of the three cases that uncertainty increases the opportunity cost of foregoing possible better experience in the bigger city (a positive θ creates an additional role for Δe). Thus, the more noisy is the signal young workers receive about their ability, the more likely they are to take the bet of locating in a big city and seeing if it fulfils their dreams.

The second difference is that now we observe some sorting by education. Workers with high and low education have different distributions of self-confidence, with the highly educated having higher values of σ on average. Since junior workers sort based on σ , there is some sorting of junior workers by education as highly educated workers tend to locate in a big city. Nevertheless, since self-confidence is a noisy signal of ability (something that workers are now aware of), sorting by education is still imperfect, so that we observe workers with both high and low education in both small and big cities.

Appendix D. Additional empirical results

Table D.1: Alternative logit estimation of the determinants of location in big and small cities

	Probability of living in big city upon completing education	For individuals living in small city upon completing education, probability of having moved to big city 10 years later	For individuals living in big city upon completing education, probability of having moved to small city 10 years later
	(1)	(2)	(3)
Self-confidence percentile	1.004 (0.002)**	1.002 (0.002)	0.998 (0.003)
Cognitive ability percentile	1.001 (0.003)	1.009 (0.004)**	1.004 (0.004)
Risk aversion percentile	1.000 (0.002)	0.993 (0.003)**	0.993 (0.003)**
Male	0.812 (0.086)**	1.014 (0.152)	1.117 (0.203)
Hispanic	2.079 (0.710)**	1.660 (0.774)	0.522 (0.216)
Black	1.389 (0.382)	1.178 (0.437)	0.377 (0.154)**
High-school graduate	0.839 (0.142)	0.772 (0.241)	0.729 (0.233)
Some college	0.888 (0.166)	1.385 (0.467)	0.767 (0.261)
College graduate	1.762 (0.492)**	1.592 (0.533)	0.774 (0.294)
Never married	1.154 (0.207)	1.062 (0.251)	0.709 (0.193)
One or more children	0.679 (0.100)***	0.734 (0.126)*	0.945 (0.206)
Full-time working spouse	1.105 (0.197)	1.200 (0.171)	0.816 (0.165)
Living in small city at age 14	0.019 (0.004)***	0.366 (0.095)***	2.449 (0.675)***
% time spent not working		1.007 (0.003)**	1.014 (0.004)***
Relative wage		0.976 (0.125)	0.953 (0.173)
Observations	3,721	2,040	1,221
Pseudo R^2	0.454	0.078	0.069

Notes: Odd ratios (exponentiated coefficients) are reported, with coefficients above one indicating a positive effect and coefficients below one indicating a negative effect. All specifications include a constant and birth-year indicators. Standard errors in parentheses are clustered at the metropolitan area level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. A 'big city' is defined as a Core Based Statistical Area (CBSA) with a population greater than 2,000,000 in 2010. White, female, ever married and high-school dropouts are the omitted categories. Relative wage is the actual pre-migration wage of the individual relative to the wage predicted for the same individual by a regression including all the other variables in column (1), plus a city-size class fixed-effect, experience in each city-size class and its square, firm tenure and its square, three-digit occupation indicators, and three-digit sector indicators.