

Multidimensional Skills, Sorting, and Human Capital Accumulation[†]

By JEREMY LISE AND FABIEN POSTEL-VINAY*

*We construct a structural model of on-the-job search in which workers differ in skills along several dimensions and sort themselves into jobs with heterogeneous skill requirements along those same dimensions. Skills are accumulated when used, and depreciate when not used. We estimate the model combining data from O*NET with the NLSY79. We use the model to shed light on the origins and costs of mismatch along heterogeneous skill dimensions. We highlight the deficiencies of relying on a unidimensional model of skill when decomposing the sources of variation in the value of lifetime output between initial conditions and career shocks. (JEL J24, J41, J64)*

The traditional approach to studying wage and employment inequality, as emphasized by Acemoglu and Autor (2011), relies on a view of labor markets where each worker is endowed with a certain level of “human capital” that rigidly dictates the type of job they are able to hold. This view has gradually evolved into one of labor markets as institutions mediating the endogenous allocation of workers with heterogeneous skills into tasks with heterogeneous skill requirements: any worker can now potentially perform any job, with their skills determining how good they are at any given job, while the market determines the assignment of skills to tasks. This more general view of labor markets has afforded great progress in our understanding of wage and employment inequality.¹

*Lise: Department of Economics, University of Minnesota, and Federal Reserve Bank of Minneapolis (email: jlise@umn.edu); Postel-Vinay: Department of Economics, University College London, and Institute for Fiscal Studies (email: f.postel-vinay@ucl.ac.uk). Luigi Pistaferri was the coeditor for this article. We would like to thank three anonymous referees plus Fatih Guvenen, Jim Heckman, John Kennan, Derek Neal, Carl Sanders, David Wiczer, Ken Wolpin and seminar audiences at ASU, Bonn, Cambridge, Chicago, Collegio Carlo Alberto (Turin), Cornell, CREST (Paris), École Polytechnique, Helsinki, Hitotsubashi University Tokyo, Lausanne, Minnesota, NYU, U Penn, Penn State, Pitt/CMU, USC, Rice, Richmond Fed, Toronto, Vienna, Washington St. Louis/St. Louis Fed and Wisconsin Madison, as well as participants in the NBER Summer Institute, RES meetings, SED meetings, SaM annual conference, Philadelphia Workshop on Macroeconomics, and Barcelona GSE Summer Forum for useful feedback on earlier versions of this paper. We are additionally grateful to Carl Sanders for kindly providing his data and helping us get started with O*NET. All errors and shortcomings are ours. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

[†]Go to <https://doi.org/10.1257/aer.20162002> to visit the article page for additional materials and author disclosure statements.

¹It shed light on secular trends such as the “polarization” of wages and employment (the simultaneous growth in wages and employment shares of both high- and low-skill workers, at the expense of the middle part of the skill distribution: Acemoglu and Autor 2011). It contributed to explaining business-cycle fluctuations in aggregate output and employment (Lise and Robin 2017). It gave substance to the intuitive notion of “skill mismatch” and helped make sense of patterns of worker turnover (Shimer and Smith 2000; Lise, Meghir, and Robin 2016). It helped clarify the informational content of wage data (Eeckhout and Kircher 2011; Hagedorn, Law, and Manovskii 2017).

A subsisting limitation of this approach, however, is that it routinely models skill and task heterogeneity as one-dimensional: workers have more or less of one catch-all “skill,” and jobs differ in their requirements for that skill. This representation is at odds with intuition: if one worker is very good at abstract problem-solving but inept at manual work, while another excels in manual tasks but struggles with abstract reasoning, how does one decide which worker is more “skilled?” It is also at odds with the perception of statistical agencies and practitioners of human resources, which maintain and use data describing workers and occupations along many different and imperfectly correlated dimensions such as years and field of education, training, health, aptitude, and psychometric test scores, etc., or the occupational skill requirements descriptors available from the O*NET program discussed below. Moreover, it is likely that workers improve the skills that they use regularly and lose some of those they do not use so much, a pattern that a scalar representation of human capital is bound to miss altogether.

The alternative view that workers are endowed with bundles of different skills used in different proportions depending on the task they perform has some history in labor economics (see Sanders and Taber 2012 for a review). But at present, few quantitative modeling tools exist that fully exploit the wealth of information on heterogeneous, multidimensional worker skills and job skill requirements available in the data in a description of labor market equilibrium.

In this paper, we contribute to filling this gap: we extend an otherwise standard and well-tested search-theoretic model of individual careers to allow for multidimensional skills and on-the-job learning. We estimate the model using occupation-level measures of skill requirements based on O*NET data, combined with a worker-level panel (NLSY79). We use the estimated model to shed light on the origins and costs of mismatch along three dimensions of skills: cognitive, manual, and interpersonal,² and the sources of variation in lifetime output.

Our main findings are the following. The model sees cognitive, manual, and interpersonal skills as very different productive attributes. Manual skills have moderate returns and adjust quickly (i.e., they are easily accumulated on the job, and relatively easily lost when left unused). Cognitive skills have much higher returns, but are much slower to adjust. Interpersonal skills have only slightly higher returns than manual skills, and are essentially fixed over a worker’s lifetime. Next, the cost of skill mismatch (modeled as the sum of an output loss and a loss of worker utility caused by skill mismatch) is very high for cognitive skills, an order of magnitude greater than for manual or interpersonal skills. Moreover, this cost is asymmetric: employing a worker who is under-qualified in cognitive skills (i.e., has a level of skills that falls short of the job’s skill requirements) is more than twice as costly in terms of lost surplus as employing an over-qualified worker. Those important differences between various skill dimensions are missed when subsuming worker productive heterogeneity into one single scalar index. Indeed, when we consider a decomposition of lifetime output, and compare our multidimensional model to a version of the model with only a single skill, the single skill version overestimates the importance of unobserved heterogeneity (relative to observed initial skills) by a

²What we term interpersonal skills are sometimes referred to as noncognitive or personality traits (Heckman, Stixrud, and Urzua 2006; Borghans et al. 2008).

factor of two, and underestimates the contribution of career shocks relative to initial observed skills by one-half.

The paper is organized as follows. Section I provides a brief discussion of some of the related literature. Section II lays out the formal model. Section III describes the data used for estimation, with some emphasis on O*NET. Section IV explains the simulation/estimation protocol. Section V presents the estimation results and discusses some of the model's predictions on skill mismatch and sorting. Section VI presents some results on the determinants of social output, and decomposes the variance of expected lifetime output into components due to initial endowments of worker skills and randomness during workers' careers. Finally, Section VII concludes.

I. Related Literature

This paper is related to the vast empirical literature on the returns to firm and occupation tenure and to more recent work on task-specific human capital (see, for example, Poletaev and Robinson 2008 and Gathmann and Schönberg 2010, among others). These literatures, and the connections between them, are covered in the excellent survey paper by Sanders and Taber (2012). As a preamble to their review of the empirical literature, Sanders and Taber (2012) offers an elegant theoretical model of job search and investment in multidimensional skills which, on many aspects, can be seen as a special case of the model in this paper.³ However, they only use their model to provide intuition and highlight key qualitative predictions of the theory, and do not bring it to the data.

In a more structural vein, attempts to model the allocation and pricing of heterogeneous supply and demand of indivisible and multidimensional bundles date back at least to Tinbergen (1956) and the hedonic model of Rosen (1974). Generic nonparametric identification of the hedonic model is established in Ekeland, Heckman, and Nesheim (2004) and Heckman, Matzkin, and Nesheim (2010). In a more recent contribution, Lindenlaub (2017) estimates the static quadratic-normal assignment model of Tinbergen (1956) along two dimensions of skills (manual and cognitive) for two different cohorts using the same combination of O*NET and NLSY data as we do in this paper. She finds an interesting pattern of technological change: the complementarity between her measures of cognitive worker skills and cognitive job skill requirements increased substantially during the 1990s, while the complementarity between manual job and worker attributes decreased. She then analyzes the consequences of that technological shift for sorting and wage inequality.

While Lindenlaub's analysis brings about many valuable new insights, it assumes away market imperfections which limits its applicability to empirical and quantitative policy analysis. First, it is difficult to define a meaningful notion of unemployment or of skill mismatch in a Walrasian (frictionless) world where, given the economy's primitives (i.e., the production technology and the distributions of job

³Sanders and Taber (2012) models individual skill accumulation as the outcome of endogenous investment decisions (in the spirit of Ben-Porath 1967), whereas we consider (occupation-specific) learning-by-doing. While the conceptual differences between those two models are important, they are notoriously difficult to tell apart empirically.

and worker attributes), equilibrium is by construction efficient. By contrast, allowing for market imperfections creates scope for welfare-improving policy intervention, the extent of the imperfections is an empirical question. Second, frictionless matching and assignment models, including Rosen's hedonic model, are static.⁴ As such, they are largely silent on questions relating to a worker's life cycle, such as the cost of skill mismatch throughout a worker's career, the way in which individual skills evolve over a career, how this skill accumulation is priced in the market, or the reasons why workers switch occupations as often as they do.⁵

Next, following in the tradition of Heckman and Sedlacek (1985), Keane and Wolpin (1997), and Lee and Wolpin (2006), the important contribution by Yamaguchi (2012) provides the first estimation of a Roy-type model of task-specific human capital accumulation and occupation choices over the life cycle based on the combination of the NLSY with data on occupation-level attributes, interpreted as "task complexity," from the Dictionary of Occupational Titles (DOT, the predecessor of O*NET). The broad approach is the same as in the present paper: each occupation is characterized by the vector of weights (the degree of task complexity) it places on a limited number of different skill dimensions, as in Lazear's (2009) skill-weights approach. Worker skills are not directly observed, but their accumulation is modeled as a hidden Markov chain, the parameters of which are identified from observed choices of occupations with different task contents (observed from the DOT data), using the model's structure. Yamaguchi's findings suggest that higher task complexity is associated with higher wage returns to, and faster growth of the skills relevant to the task. A wage variance decomposition further suggests that both cognitive and motor skills (the two skill dimensions considered by Yamaguchi) are important determinants of cross-sectional log wage variance. A decomposition of wage growth shows that cognitive skills account for all of the wage growth of high-school and college graduates, while motor skills account for about one-half of the wage growth of high-school dropouts.

While clearly related in spirit, our model differs from Yamaguchi (2012) in several important ways. First, Yamaguchi (2012) is a frictionless model in which occupation mobility is largely governed by unobserved shocks to an exogenously posited wage function, unobserved shocks to workers' skills, and unobserved shocks to workers' preferences for any given type of job.⁶ We propose a more parsimonious

⁴See Chiappori and Salanié (2016) for a recent survey of the econometrics of static, frictionless matching models.

⁵Models of experimentation and learning following on from Jovanovic (1979) such as Neal (1999), Pavan (2011), and Antonovics and Golan (2012) have been useful to make sense of the patterns of between and within occupation switches in the NLSY. They cannot, however, easily rationalize the frequent transitions in and out of unemployment observed in the same data. Lindenlaub and Postel-Vinay (2020) extends Lindenlaub's frictionless model to a frictional environment using a basic framework that has much in common with our model. However, Lindenlaub and Postel-Vinay (2020) is a theoretical exercise focusing on conditions under which specific sorting patterns emerge in steady-state equilibrium, and their model does not feature human capital accumulation. Anderson and Smith (2010) develops a matching model with evolving one-dimensional types in a frictionless environment. They characterize the conditions for assortative matching in this context, as well as implications for life-cycle dynamics of wages.

⁶One immediately apparent drawback of this frictionless approach is that, taken literally, it predicts that workers should change occupations continuously (or in every period, in Yamaguchi's discrete-time model), which is obviously at odds with observation. Note that any dynamic model that aims to replicate the empirical job mobility patterns will need to incorporate some type of mobility friction (for example, a mobility cost) to prevent workers from continuously reallocating in the face of new shocks or new information. For example, Guvenen et al. (2020),

random search model, in which the only shocks are the receipt of job offers by workers. Wages and mobility decisions are then endogenously determined through between-employer competition for labor services. Our less flexible, but more transparent and readily interpretable specification offers a remarkably good fit to the data. Second, the only engine of wage growth in Yamaguchi's model is skill accumulation. Other sources of wage growth, such as job-shopping or learning, are therefore partly picked up by skills in that model, which may lead to an overstatement of the role of skills. Our model also ignores learning, but explicitly models job-shopping as an additional source of wage dynamics.⁷ Adding search frictions allows us to address issues related to unemployment and skill mismatch.⁸

Two recent papers are particularly related. Taber and Vejlin (2020) estimates a model which allows for search, Roy-type selection, human capital accumulation, and non-wage amenities. Workers are modeled as having a time invariant relative ability at each job-type in the economy. In the absence of frictions they would choose a single job-type and remain indefinitely. Human capital is assumed to be general and accumulated while working. Job mobility is informative about the degree of search frictions, and wage cuts are informative about non-wage amenities. Taber and Vejlin (2020) models relative ability between jobs/occupations as an unobserved vector with dimension equal to the number of job-types in the economy. We take a substantially more parsimonious approach: a worker's relative productivity across jobs/occupations depends on the amount of skills (cognitive, manual, interpersonal ...) they currently possess and whether they are a good fit for the demands of a particular job. Rather than treating these skills as unobserved, we use a large set of premarket measurements to estimate a worker's initial endowment, and a similar large set of measurements on occupations to estimate the skill requirements of jobs. A second notable modeling difference is that we allow these skills to evolve differentially depending on the extent to which they are being used in a particular job/occupation. We are particularly interested in the differential returns to these skills, and the extent to which each type of skill can be learned on the job.

From an empirical perspective, perhaps the closest paper is Guvenen et al. (2020). They use the same combination of NLSY and O*NET data as we do to construct a summary index of multidimensional skill mismatch which they use to assess the impact of skill mismatch on wages and patterns of occupational switching. They produce a rich set of empirical results, a rough summary of which is that both current and past mismatch strongly impact wages (negatively), the probability of switching occupations (positively), and the direction of said switching.

Their index of skill mismatch is derived from a model of occupation choice with workers learning about their own ability. Mismatch arises in this model, not because of search frictions, but because workers have imperfect information about their own skills and sort into occupations that are optimal for their *perceived* skill bundle (which differs from their true one). As they gradually learn about their true skills

which we discuss below, relies on a combination of information and mobility frictions to generate the empirical implications of their learning model.

⁷Sanders (2012) considers learning in a model otherwise similar to Yamaguchi's.

⁸Moscarini (2001) combines a two-sector Roy model into an equilibrium matching model and analyzes the partially directed search patterns arising in equilibrium and governing equilibrium selection of skill bundles into sectors. His setup has great descriptive appeal, but remains far too stylized to be taken directly to the data.

(about which they observe a sequence of noisy signals over time), workers switch occupations. This model gives rise to an intuitive summary mismatch measure that is based on the distance between a worker's skill bundle and the set of skills required by their occupation, which the authors then use as a regressor in Mincer-type wage equations and in statistical models of occupation switching.

While our paper shares some of its basic objectives with Guvenen et al. (2020) (chiefly, an assessment of the production/wage cost of skill mismatch in various dimensions), the two contributions differ in terms of both approach and focus. Aside from substantive differences in modeling choices, Guvenen et al. use their theory as a guide for intuition and specification of reduced-form statistical models rather than as an actual structure for estimation. More importantly, they provide detailed results on the impact of mismatch on the probability and direction of occupational switching, whereas we focus (i) on differences between skill categories in the speed of human capital accumulation or decay and (ii) on the social cost of various forms of mismatch. Our structural approach is especially useful to address the latter broad question, which we do by means of counterfactual simulations.

II. Job Search with Multidimensional Job and Worker Attributes

A. The Model

The Environment.—Workers are characterized by general and specialized skills. The market productivity of specialized skills depends on the technology of a particular firm, while general skills have a common effect on output, independent of the particular firm technology a worker is currently matched with. Match output is $f(\mathbf{x}, \mathbf{y})$, where $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^K$ describes the worker's set of skills, and $\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^L$ describes the firm's technology, with $L \leq K$. The first L worker skills are specialized with the remaining $K - L$ being general skills. Time is continuous. The firm's technology is fixed, but the worker's skills gradually adjust to the firm's technology as follows:

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, \mathbf{y}),$$

where $\mathbf{g}: \mathbb{R}^K \times \mathbb{R}^L \rightarrow \mathbb{R}^K$ is a continuous function. Just as in production, the adjustment of specialized skills differs depending on the firm technology, while the adjustment of general skills depends only on experience.

Upon entering the labor market, workers draw their initial skill vectors from an exogenous distribution $N(\mathbf{x})$, with density $\nu(\mathbf{x})$. Workers can be matched to a firm or unemployed. If matched, they lose their job at rate δ , and they sample alternate job offers from the fixed sampling distribution $\Upsilon(\mathbf{y})$, with density $v(\mathbf{y})$, at rate λ_1 . Unemployed workers sample job offers from the same sampling distribution at rate λ_0 . Workers exit the market at rate μ . All four transition rates $(\lambda_0, \lambda_1, \delta, \mu)$ are exogenous.

All agents have linear preferences over income and discount the future at rate r . A type- \mathbf{x} worker's flow utility from working in a type- \mathbf{y} job for a wage w is $w - c(\mathbf{x}, \mathbf{y})$, where $c(\mathbf{x}, \mathbf{y})$ is disutility from work, which depends on the type of the match, (\mathbf{x}, \mathbf{y}) .

A type- \mathbf{x} unemployed worker receives a flow income $b(\mathbf{x})$ and has no disutility of being unemployed.

Firm, Worker, and Match Values.—We denote the total private value (i.e., the value to the firm-worker pair) of a match between a type- \mathbf{x} worker and a type- \mathbf{y} firm by $P(\mathbf{x}, \mathbf{y})$. Under linear preferences over wages, this value is independent of the way in which it is shared between the two parties, and only depends on match attributes (\mathbf{x}, \mathbf{y}) . We further denote the value of unemployment by $U(\mathbf{x})$, and the worker's value of their current wage contract by W , which we discuss in detail below. Admissible worker values imply $W \geq U(\mathbf{x})$ (otherwise the worker would quit into unemployment), and $W \leq P(\mathbf{x}, \mathbf{y})$ (otherwise the firm would fire the worker). Assuming that the employer's value of a job vacancy is zero (which would arise under free entry and exit of vacancies on the search market), the total *surplus* generated by a type- (\mathbf{x}, \mathbf{y}) match is $P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})$, and the worker's share of that surplus is $(W - U(\mathbf{x})) / (P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}))$.

Rent Sharing and Wages.—Wage contracts are renegotiated sequentially by mutual agreement, as in the sequential auction model of Postel-Vinay and Robin (2002). Workers have the possibility of playing off their current employer against any firm from which they receive an outside offer. If they do so, the current and outside employers Bertrand-compete over the worker's services.⁹

Consider a type- \mathbf{x} worker employed at a type- \mathbf{y} firm and assume that the worker receives an outside offer from a firm of type \mathbf{y}' . Bertrand competition between the type- \mathbf{y} and type- \mathbf{y}' employers implies that the worker ends up in the match that has higher total value, that is, they stay in their initial job if $P(\mathbf{x}, \mathbf{y}) \geq P(\mathbf{x}, \mathbf{y}')$ and moves to the type- \mathbf{y}' job otherwise, with a new wage contract worth $W' = \min\{P(\mathbf{x}, \mathbf{y}), \max\{P(\mathbf{x}, \mathbf{y}'), W\}\}$.

Suppose, for the sake of argument, that $P(\mathbf{x}, \mathbf{y}) \geq P(\mathbf{x}, \mathbf{y}') > W$. In this case, the outcome of the renegotiation is such that the worker stays with their initial type- \mathbf{y} employer under a new contract with value $W' = P(\mathbf{x}, \mathbf{y}')$. The worker's renegotiated share of the match surplus, denoted $\sigma(\mathbf{x}, \mathbf{y}, \mathbf{y}')$, is therefore

$$(1) \quad \sigma(\mathbf{x}, \mathbf{y}, \mathbf{y}') = \frac{P(\mathbf{x}, \mathbf{y}') - U(\mathbf{x})}{P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})} \in [0, 1].$$

The value $W' = P(\mathbf{x}, \mathbf{y}') = U(\mathbf{x}) + \sigma(\mathbf{x}, \mathbf{y}, \mathbf{y}') [P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})]$ may be delivered over time to the worker in a multiple ways. To pin this down we assume that the surplus share σ , negotiated at the time the worker receives an outside offer from the type- \mathbf{y}' job, stays constant until the following renegotiation. Put differently, while the worker's skill bundle \mathbf{x} and, as a consequence, the match surplus $P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})$ evolve over the course of their tenure in the type- \mathbf{y} job, the share of that surplus transferred to the worker stays constant between negotiations and is determined as per equation (1) by the best outside offer previously received by the worker. The

⁹Obviously, renegotiation only takes place if $P(\mathbf{x}, \mathbf{y}') > W$, as otherwise the type- \mathbf{y}' employer is unable to make a (profitable) offer that improves on the worker's initial value W , and the worker's threat of accepting an offer from that employer is not credible.

particular way in which the type- \mathbf{y} employer delivers the value $P(\mathbf{x}, \mathbf{y}')$ to the worker only affects the time profile of wage payments and the timing of renegotiation. It makes no difference to the allocation of workers into jobs, as mobility decisions are only based on comparisons of total match values, which, under linear preferences, are independent of the time profile of wage payments.¹⁰

Value Functions and Wage Equation.—The total private value of a match between a type- \mathbf{x} worker and a type- \mathbf{y} firm, $P(\mathbf{x}, \mathbf{y})$, solves¹¹

$$\begin{aligned} (2) \quad (r + \mu + \delta)P(\mathbf{x}, \mathbf{y}) &= f(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) + \delta U(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x}, \mathbf{y}) \\ &\quad + \lambda_1 E[W(\mathbf{x}, \mathbf{y}', \sigma) - P(\mathbf{x}, \mathbf{y}) | P(\mathbf{x}, \mathbf{y}') > P(\mathbf{x}, \mathbf{y})] \\ &= f(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) + \delta U(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Note that Bertrand competition implies that the frequency at which the worker collects offers, λ_1 , does not affect $P(\mathbf{x}, \mathbf{y})$. Upon receiving an outside offer, the worker either stays in his initial match, in which case the continuation value for that match is $P(\mathbf{x}, \mathbf{y})$, or they accept the offer, in which case they extract a value of $P(\mathbf{x}, \mathbf{y})$ from the poacher (as a result of Bertrand competition) and leave their initial employer with a vacant job worth 0. Either way, the joint continuation value for the partners in the initial match equals $P(\mathbf{x}, \mathbf{y})$. This is a key implication of Bertrand competition between employers. From a social perspective, cases where the worker accepts the outside offer and moves to a match with higher value $P(\mathbf{x}, \mathbf{y}')$ are associated with a net surplus gain of $P(\mathbf{x}, \mathbf{y}') - P(\mathbf{x}, \mathbf{y})$. Yet none of the social gains associated with future job mobility are internalized by private agents, as those gains accrue to a third party (the worker's future employer).¹²

The value of unemployment, $U(\mathbf{x})$, solves

$$(3) \quad (r + \mu)U(\mathbf{x}) = b(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \mathbf{0}) \cdot \nabla U(\mathbf{x}),$$

where the employer type is set to $\mathbf{y} = \mathbf{0}_L$ for an unemployed worker. For reasons similar to those just discussed about $P(\mathbf{x}, \mathbf{y})$, the worker fails to internalize the gain in surplus associated with accepting a job offer, and the private value of unemployment is independent of the frequency at which those offers arrive.

¹⁰Common alternative assumptions about the way in which firms deliver value to workers include a constant wage or a constant share of match output (a piece rate). Our assumption of a constant surplus share has the merit of simplifying computations considerably. Note that the wages produced by the constant wage, constant piece rate, or constant surplus share assumptions are exactly identical if the worker's skills stay constant over time ($\dot{\mathbf{x}} \equiv 0$).

¹¹The dot ("·") denotes the outer product, ∇ denotes the gradient, and $\nabla_{\mathbf{x}}$ denotes the gradient with respect to the subset \mathbf{x} of the function's arguments.

¹²We discuss some of the consequences of this property in Section VD and in online Appendix Section B.B1.

The worker receives an endogenous share σ of the match surplus $P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})$, which they value at $W(\mathbf{x}, \mathbf{y}, \sigma) = (1 - \sigma)U(\mathbf{x}) + \sigma P(\mathbf{x}, \mathbf{y})$. The wage $w(\mathbf{x}, \mathbf{y}, \sigma)$ implementing that value solves

$$(4) \quad (r + \delta + \mu)W(\mathbf{x}, \mathbf{y}, \sigma) = w(\mathbf{x}, \mathbf{y}, \sigma) - c(\mathbf{x}, \mathbf{y}) + \delta U(\mathbf{x}) \\ + \lambda_1 E \max \{0, \min \{P(\mathbf{x}, \mathbf{y}), P(\mathbf{x}, \mathbf{y}')\} - W(\mathbf{x}, \mathbf{y}, \sigma)\} \\ + \mathbf{g}(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} W(\mathbf{x}, \mathbf{y}, \sigma),$$

where the expectation is taken over the sampling distribution, $\mathbf{y}' \sim \Upsilon$.

Combining (2), (3), and (4) (using $W(\mathbf{x}, \mathbf{y}, \sigma) = (1 - \sigma)U(\mathbf{x}) + \sigma P(\mathbf{x}, \mathbf{y})$) yields the following wage equation:

$$(5) \quad w(\mathbf{x}, \mathbf{y}, \sigma) = \sigma f(\mathbf{x}, \mathbf{y}) + (1 - \sigma)[b(\mathbf{x}) + c(\mathbf{x}, \mathbf{y})] \\ - \lambda_1 E \left[\max \{0, \min \{P(\mathbf{x}, \mathbf{y}') - P(\mathbf{x}, \mathbf{y}), 0\} \right. \\ \left. + (1 - \sigma)(P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})) \right] \\ - (1 - \sigma)(\mathbf{g}(\mathbf{x}, \mathbf{y}) - \mathbf{g}(\mathbf{x}, \mathbf{0})) \cdot \nabla U(\mathbf{x}).$$

The first term $\sigma f(\mathbf{x}, \mathbf{y}) + (1 - \sigma)[b(\mathbf{x}) + c(\mathbf{x}, \mathbf{y})]$ reflects static sharing of the match surplus flow, in shares $(\sigma, 1 - \sigma)$ resulting from the worker's history of outside job offers. Note that the worker always has to be compensated for a share $(1 - \sigma)$ of their foregone home production and disutility of work $c(\mathbf{x}, \mathbf{y})$. The next (expectation) term reflects value of future outside offers, which the worker pays for by accepting a lower starting wage. The final term reflects the fact that an employed worker's skill bundle evolves toward the job's skill requirements \mathbf{y} , whereas those skills would erode toward $\mathbf{0}_L$ if the worker was unemployed. This, in general, benefits the worker in the event they become unemployed, and therefore affects the wage negatively.¹³

B. Model Analysis

A Fully Closed-Form Case.—Full closed-form solutions can be obtained under specific functional form assumptions. We now give an example, which we will use in our empirical specification below.

We first restrict the dimensionality of worker and job attributes, both for simplicity of exposition and because those restrictions are relevant to the empirical application below (nothing in the theory depends on those particular restrictions). We think of a typical worker's skill bundle $\mathbf{x} = (x_C, x_M, x_I, x_T)$ as capturing (i) the worker's cognitive skills x_C , (ii) the worker's manual skills x_M , (iii) the worker's interpersonal

¹³ As mentioned in Section I, the model in Sanders and Taber (2012) is close to a special case of our model where $f(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ and where workers always receive a fixed share of the match surplus (i.e., σ is a fixed constant). Predicted wages differ between our model and theirs, but the worker-job allocation for given distributions of \mathbf{x} and \mathbf{y} is identical. As already mentioned, the two models further differ in the specific assumption regarding skill accumulation (endogenous investment decisions versus learning-by-doing).

skills x_I , and (iv) the worker's "general efficiency" x_T . Jobs are likewise characterized by a three-dimensional bundle $\mathbf{y} = (y_C, y_M, y_I)$ capturing measures of the job's requirements in cognitive, manual, and interpersonal skills. All three job attributes are fixed over time, whereas a worker's cognitive, manual, and interpersonal skills (x_C, x_M, x_I) are allowed to adjust over time to the requirements of the particular job the worker holds (learning by doing).

The key functional form assumption is to assume a linear adjustment for skills. In particular, we assume that a worker's specialized (i.e., cognitive, manual, and interpersonal) skills adjust linearly to their job's skill requirements:

$$(6) \quad \mathbf{g}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \dot{x}_C \\ \dot{x}_M \\ \dot{x}_I \\ \dot{x}_T \end{pmatrix} = \begin{pmatrix} \gamma_C^u \max\{y_C - x_C, 0\} + \gamma_C^o \min\{y_C - x_C, 0\} \\ \gamma_M^u \max\{y_M - x_M, 0\} + \gamma_M^o \min\{y_M - x_M, 0\} \\ \gamma_I^u \max\{y_I - x_I, 0\} + \gamma_I^o \min\{y_I - x_I, 0\} \\ g_{x_T} \end{pmatrix},$$

where the $\gamma_k^{u/o}$ terms are all positive constants governing the speed at which worker skills adjust to a job's requirements. Note that we allow that speed to differ between upward and downward adjustments (γ_k^u versus γ_k^o for $k = C, M, I$, where u stands for "under-qualified" and o stands for "over-qualified"), and between skill types ($\gamma_C^{u/o}$ versus $\gamma_M^{u/o}$ versus $\gamma_I^{u/o}$). In this case a worker's skills relate to job tenure $s - t$ as follows:

$$(7) \quad x_k(s) = y_k - e^{-\gamma_k^{u/o}(s-t)}(y_k - x_k(t)),$$

where the adjustment speed $\gamma_k^{u/o}$ that applies depends on whether $k = C, M$ or I and whether $x_k(t) \geq y_k$ (see Appendix Section AA for details). Over time a worker's specialized skills will adjust to the requirements of the job. Finally, a worker's general efficiency simply grows at a constant rate: $x_T(t) = x_T(0) \times e^{g_T t}$, independently of the worker's cognitive/manual skills or of the worker's employment status. This simple specification will help the model capture the wage/experience trend observed in the data.

The production function is then specified as follows:

$$(8) \quad f(\mathbf{x}, \mathbf{y}) = x_T \times \left[\alpha_T + \sum_{k=C, M, I} (\alpha_k y_k - \kappa_k^u \min\{x_k - y_k, 0\}^2 + \alpha_{kk} x_k y_k) \right].$$

In this specification of $f(\mathbf{x}, \mathbf{y})$, the linear terms $\alpha_k y_k$ capture the possibility that jobs with different requirements in any of cognitive, manual, or interpersonal skills have inherently different productivity levels, regardless of the worker they are matched with. The next three terms, $-\kappa_k^u \min\{x_k - y_k, 0\}^2$, $k = C, M, I$ capture the idea that a worker with a shortage of skills \mathbf{x} compared to the job's skill requirement level \mathbf{y} in any dimension (cognitive, manual, or interpersonal) causes a loss of output (assuming that all κ_k^u values are nonnegative). We allow for the output loss caused by skill mismatch to differ depending on which skills the worker is deficient in. The last three terms, $\alpha_{kk} x_k y_k$, $k = C, M, I$, are added to provide additional flexibility in the modeling of complementarity between worker skills and job skill requirements within all three skill categories. In particular, assuming that $\alpha_{kk} \geq 0$, those last terms imply that an over-qualified worker (with $x_k > y_k$) produces more output

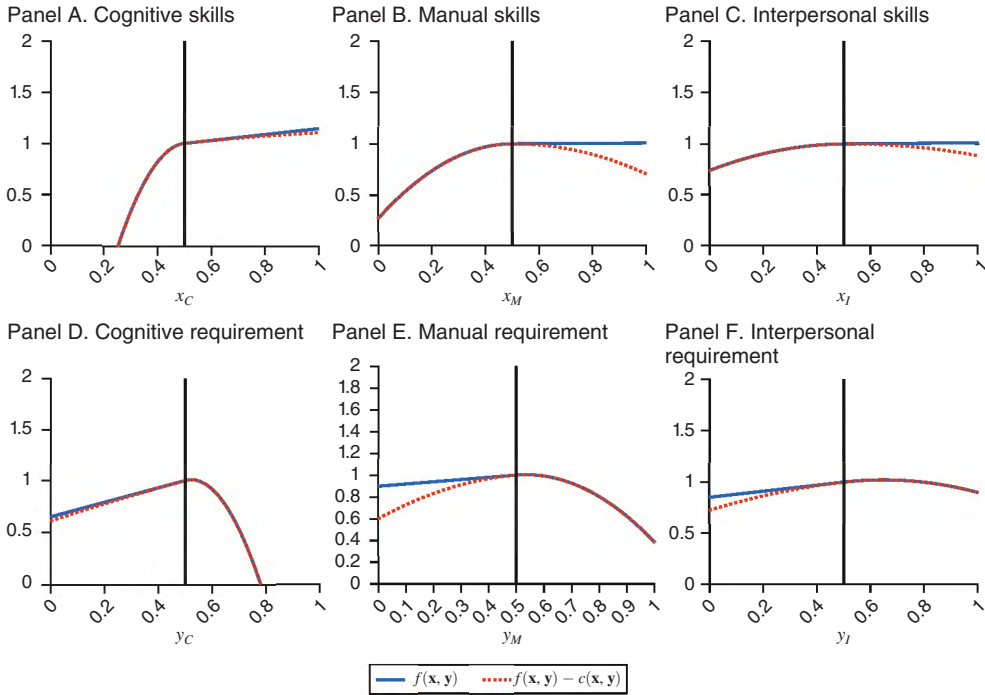


FIGURE 1. THE PRODUCTION FUNCTION

than a worker with $x_k = y_k$, whose skills are an exact match for the job's requirements in skill dimension k . Finally, general efficiency, x_T , merely scales output up or down, conditional on \mathbf{x} and \mathbf{y} . In the estimation we will allow x_T to be correlated with initial skill bundles \mathbf{x}_0 , allowing for the possibility that workers with, for example, high cognitive skills also have high general skills.

Figure 1 provides a visual impression of the way in which the production function $f(\mathbf{x}, \mathbf{y})$ varies with its various inputs. The solid lines on the six panels show the response of $f(\mathbf{x}, \mathbf{y})$ to variations in each of the three components of \mathbf{x} (top row) and of \mathbf{y} (bottom row), holding all other components of \mathbf{x} and \mathbf{y} fixed at 0.5. All graphs are put on a common scale, with $f(\mathbf{x}, \mathbf{y})$ normalized to 1 at the point where $\mathbf{x} = \mathbf{y}$, and constructed using our parameter estimates, presented below.

The next object we need to specify is the flow disutility of work:

$$(9) \quad c(\mathbf{x}, \mathbf{y}) = x_T \times \sum_{k=C,M,I} \kappa_k^o \max\{x_k - y_k, 0\}^2.$$

According to this specification, disutility of work is only positive if the worker is over-qualified for their job in some skill dimension. We interpret this as a utility cost of being under-matched. Note that our specification of $c(\mathbf{x}, \mathbf{y})$ is formally identical to the one we use to capture the output loss caused by a *shortage* of worker cognitive, manual, or interpersonal skills compared to the job's requirements in the production function (8). The utility cost of being under-matched further allows for an *excess* of skills to cause a loss of match value, albeit without causing a loss of output. An appealing implication of this specification is that over-qualified workers will have to

be compensated for that utility cost, and will therefore have to be paid more in a given job than workers whose skills exactly match the job's requirements. Again, a visual impression of the flow surplus from a match, $f(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y})$, is given by the dotted lines on Figure 1, which are constructed in the same way as the corresponding production function lines discussed above. The relatively small vertical distance between the solid and dotted lines in regions of the graphs where the worker is over-qualified suggests that the utility cost of over-qualification is quantitatively relatively small, a point to which we will return when we discuss estimation results.

Finally, for simplicity, we specify unemployment income as depending on general skill only, $b(\mathbf{x}) = bx_T$, with b a positive constant, so that $U(\mathbf{x}) = bx_T/(r + \mu - g)$ is independent of the specialized skills (x_C, x_M, x_I) .

With those specifications, equations (2) and (3) imply (see Appendix Section AA):

$$\begin{aligned}
 (10) \quad & P(\mathbf{x}(t), \mathbf{y}) - U(\mathbf{x}(t)) \\
 &= x_T(t) \times \left\{ \frac{\alpha_T + \sum_{k=C,M,I} (\alpha_k y_k + \alpha_{kk} y_k^2) - b}{r + \delta + \mu - g} \right. \\
 &\quad - \sum_{k=C,M,I} \left(\frac{\kappa_k^u \min\{x_k(t) - y_k, 0\}^2}{r + \delta + \mu - g + 2\gamma_k^u} + \frac{\kappa_k^o \max\{x_k(t) - y_k, 0\}^2}{r + \delta + \mu - g + 2\gamma_k^o} \right) \\
 &\quad \left. + \sum_{k=C,M,I} \alpha_{kk} y_k \times \left(\frac{\min\{x_k(t) - y_k, 0\}}{r + \delta + \mu - g + \gamma_k^u} + \frac{\max\{x_k(t) - y_k, 0\}}{r + \delta + \mu - g + \gamma_k^o} \right) \right\}.
 \end{aligned}$$

The first term in the equation defining match surplus $P(\mathbf{x}(t), \mathbf{y}) - U(\mathbf{x}(t))$ is the surplus achieved if the worker's skills are perfectly matched to the job's requirements, i.e., if $(x_C(t), x_M(t), x_I(t)) = (y_C, y_M, y_I)$. The remaining terms reflect the net private surplus cost of initial cognitive, manual, and interpersonal skill mismatch. This cost obviously depends on the weights of cognitive, manual, and interpersonal mismatch in the technology and utility function, and on their degree of complementarity between worker skills and job skill requirements. But it also depends on the speed of skill adjustment: if adjustment is instantaneous ($\gamma_k^{u/o} \rightarrow +\infty$), the cost of mismatch becomes negligible.

III. Data

Our estimation sample is a panel of worker-level data from the 1979 National Longitudinal Survey of Youth (NLSY79, Bureau of Labor Statistics 1979–2011) combined with occupation-level data on skill requirements from the O*NET program (National Center for O*NET Development 2014). We describe both datasets and the way we combine them before turning to a description of the estimation sample itself.¹⁴ Additional details can be found in Appendix Section AB.

¹⁴We are not the first authors to combine these data sources. A non-exhaustive list includes Autor, Levy, and Murnane (2003); Acemoglu and Autor (2011); Autor and Dorn (2013b); Yamaguchi (2012); Sanders (2012);

A. Construction of the Estimation Sample

Data Sources.—The NLSY79 is well known and requires little description. Our extract from that dataset is a weekly unbalanced panel of workers whom we follow from first entry into the labor market. For each worker in the panel, time is set to zero at the first week they cease to be in full-time education. We focus on males from the main sample who were never in the military,¹⁵ and retain all individual histories until the first occurrence of a non-employment spell of 18 months or more. We consider individuals experiencing such a long spell of non-employment as losing their attachment to the labor force, which we treat as attrition from the sample. We retain information on labor force status and transitions, weekly earnings, occupation of current job (Census codes), education (highest grade completed), performance in a battery of ten aptitude tests called the Armed Services Vocational Aptitude Battery (ASVAB), measures of antisocial behavior, measures of health, and scores in two psychometric tests measuring social skills. Education, ASVAB scores, and measures of social skills and health will be used as measures of the initial skill bundles \mathbf{x}_0 of those workers (more below).

To obtain measures of the skill requirements \mathbf{y} attached to the occupations observed in the NLSY sample, we combine the latter with data from the O*NET program. O*NET (the Occupational Information Network) is a database describing occupations in terms of skill and knowledge requirements, work practices, and work settings.¹⁶ It comes as a list of 277 descriptors, with ratings of importance, level, relevance or extent, for over 970 different occupations. O*NET descriptors are organized into nine broad categories: skills, abilities, knowledge, work activities, work context, experience/education levels required, job interests, work values, and work styles. O*NET ratings come from two different sources: a survey of workers, who are asked to rate their own occupation in terms of a subset of the O*NET descriptors, and a survey of “occupation analysts” who are asked to rate other descriptors in the O*NET set.

We retain descriptors from the skills, abilities, knowledge, work activities, and work context O*NET files, as descriptors contained in the other files (job interests, work values, and work styles) are less directly interpretable in terms of skill requirements, and merge those files with our NLSY sample, based on occupation codes.^{17,18}

Lindenlaub (2017); and Guvenen et al. (2020), who all use combinations of the NLSY with occupation data from O*NET or from its predecessor, the Dictionary of Occupational Titles.

¹⁵The NLSY over-samples ethnic minorities, people in the military, and the poor. We drop all such over-sampled observations.

¹⁶O*NET is developed by the North Carolina Department of Commerce and sponsored by the US Department of Labor. Its initial purpose was to replace the old Dictionary of Occupational Titles. More information is available at www.onetcenter.org, or the related Department of Labor site (www.doleta.gov/programs/onet/eta_default.cfm).

¹⁷The NLSY79 uses 1970, 1980, and 2000 Census codes for occupation, whereas O*NET uses 2009 SOC codes. Crosswalks exists between those different nomenclatures. The crosswalks we use were kindly provided to us by Carl Sanders, whose help is gratefully acknowledged. Using those, over 92 percent of occupation codes records in the NLSY sample have a match in the O*NET data.

¹⁸In our data sample we define a job change as a change of employer. In some cases, individuals are observed changing occupations *within an employer spell*. The typical pattern in those cases is that workers are oscillating between two or three “similar-looking” occupations (for example, an individual in our sample has a spell of employment starting with 10 weeks as an Electrical Engineer, continuing with 132 weeks as an Aerospace Engineer, followed by 50 weeks as a Mechanical Engineer, and ending with 9 weeks as an Aerospace Engineer again). This

Job Skill Requirements.—Our selection from the O*NET database leaves us with over 200 different descriptors, which we take as measures of the underlying skill requirements. We reduce this large set of descriptors to three dimensions, which we interpret as “cognitive,” “manual,” and “interpersonal” skill requirements, using the following procedure.¹⁹ First, we run Principal Component Analysis (PCA) on our large set of O*NET measures and keep the first three principal components. We then recover our cognitive, manual, and interpersonal skill requirement indices by recombining those three principal components (which by default are constructed to be orthonormal) in such a way that they satisfy the following three exclusion restrictions: (i) the *mathematics* score only reflects cognitive skill requirements; (ii) the *mechanical knowledge* score only reflects manual skill requirements; (iii) the *social perceptiveness* score only reflects interpersonal skill requirements. Interpretation of the three skill requirement indices thus obtained as cognitive, manual, and interpersonal therefore relies on those exclusion restrictions. Finally, we rescale our skill requirement indices so that they lie in $[0, 1]$.²⁰

Worker Skill Bundles.—Finally, we need to construct a distribution of initial worker skill bundles, i.e., the distribution $N(\mathbf{x})$ of cognitive, interpersonal, and manual skills among labor market entrants. For this we follow a similar procedure as for the distribution of skill requirements, using PCA and exclusion restrictions. We use the following sets of measures: the ten ASVAB scores that are directly available from the NLSY sample; individual scores on the Rotter locus-of-control scale and the Rosenberg self-esteem scale tests;²¹ three measures of criminal and antisocial behavior; two measures of health (BMI and weight); and an O*NET-based measure of cognitive, manual, and interpersonal skills attached to the level of education attained by each NLSY sample member. The latter is constructed using the “experience/education requirements” file from O*NET, which informs about the education requirements of each occupation in O*NET, and from which we take the average value, for each education level, of the cognitive, manual, and interpersonal scores constructed above. As exclusion restriction we assume that (i) the ASVAB *mathematics knowledge* score only reflects cognitive skills; (ii) the ASVAB *automotive and shop information* score only reflects manual skills; (iii) the Rosenberg *self-esteem* score only reflects interpersonal skills. Those particular exclusion restrictions were chosen for their intuitive consistency with the exclusion restrictions used in the construction of job skill requirements, so as to ensure that worker skill indices are reasonably well “aligned” with the corresponding skill requirement indices in all three dimensions. Yet in the estimation, we will allow for the possibility of

is strongly suggestive of classification error. To address that, we average the skill requirements of those occupations within an employer spell.

¹⁹ Technical details of the construction of our job skill requirement and worker skill bundles are given in Appendix Section AC.

²⁰ We do this using linear transforms (rather than by converting the initial indices to ranks, as has been done elsewhere in the literature), because we expect there to be useful information in the distance between two different occupations in terms of cognitive, manual, or interpersonal skill requirements. Linear rescaling preserves relative distances, whereas conversion into ranks renders all adjacent occupations equidistant.

²¹ See <https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/attitudes> for a description of those two tests.

less-than-perfect alignment between worker skill and job skill requirement scores (see below).

B. Empirical Content of Skill and Skill Requirement Bundles

In this subsection we present reduced-form evidence that our measures of worker skills and job skill requirements have predictive power for wages in the current job. We also present evidence that our measure of a worker's skills, deviation from their job's skill requirements, and duration in their first job have predictive power for the skill requirements of the job they move to next. We also compare our measure to some common alternatives from the literature.

Skills, Skill Requirements, and Wages.—Our measures of occupational skill requirements and initial worker skill bundles are not particularly sensitive to our choice of exclusion restrictions (see Appendix Section AD for details on robustness). Additionally, because our skill requirement measures make maximal use of the variation contained in the O*NET, they out perform the abstract-, manual-, and routine-task measures proposed by Autor and Dorn (2013b)—henceforth, AD—in terms of explanatory power in a descriptive wage regression. This can be seen in Table 1, which presents the results from regressions of log wages on various sets of skill and skill requirement measures. The regressors include our measures of initial cognitive, manual, and interpersonal skills (x_{C0} , x_{M0} , x_{I0}), our measures of cognitive-, manual-, and interpersonal-skill requirements for the current occupation (\tilde{y}_C , \tilde{y}_M , \tilde{y}_I), and the AD measures of the abstract-, manual-, and routine-task intensity of the current occupation ($\tau_{abstract}^{AD}$, τ_{manual}^{AD} , $\tau_{routine}^{AD}$). The adjusted R^2 when including our skill requirement measures is higher than when including the AD task measures (0.37 compared to 0.34). When both sets of measures are included in the regression, the adjusted R^2 barely increases relative to the specification using only our measures, and many of the coefficient estimates on the AD task measures become small and statistically insignificant, while our measures continue to have the same statistically significant coefficients as when entered alone. These results hold both when we include only level effects and interactions between initial worker skills and job skill requirement (columns 1–3 compared to 4–6). We interpret the results from this descriptive regression as indicating that our measures contain predictive power for wages and that they contain all the information in the AD task measures, and more.²²

In our model we represent an occupation by a vector of three skill requirements: cognitive, manual, and interpersonal. This is a parsimonious representation with a natural measure of how similar occupations are to each other. We represent how good an individual is at an occupation by how close their own skill vector is to the occupation skill vector. This representation provides, for each worker, a complete ranking of how good they are at each occupation, and assumes they will be similarly good at occupations that have similar skill requirements. An alternative would be to directly represent how good each worker is at each occupation, without imposing

²² We are able to create consistent occupational classifications for both our measures and the AD measures for 97 percent of our observations.

TABLE 1—EMPIRICAL CONTENT OF SKILL AND TASK MEASURES

log wage	(1)	(2)	(3)	(4)	(5)	(6)
x_{C0}	0.339 (0.115)	0.483 (0.117)	0.316 (0.115)	−0.054 (0.152)	0.318 (0.129)	−0.076 (0.153)
x_{M0}	−0.088 (0.087)	−0.093 (0.089)	−0.070 (0.087)	0.064 (0.168)	−0.122 (0.099)	0.024 (0.169)
x_{J0}	0.268 (0.053)	0.311 (0.054)	0.262 (0.053)	0.234 (0.101)	0.394 (0.095)	0.305 (0.141)
\tilde{y}_C	0.657 (0.071)		0.704 (0.082)	0.048 (0.165)		0.197 (0.185)
\tilde{y}_M	0.259 (0.058)		0.170 (0.066)	0.418 (0.165)		0.402 (0.178)
\tilde{y}_I	0.389 (0.063)		0.285 (0.066)	0.411 (0.142)		0.361 (0.147)
$\tau_{abstract}^{AD}$		0.483 (0.032)	0.121 (0.034)		0.191 (0.104)	−0.052 (0.115)
τ_{manual}^{AD}		0.251 (0.048)	0.194 (0.050)		−0.017 (0.177)	−0.291 (0.179)
$\tau_{routine}^{AD}$		0.171 (0.028)	−0.012 (0.030)		0.273 (0.088)	0.071 (0.089)
$x_{C0} \times \tilde{y}_C$				0.902 (0.221)		0.691 (0.250)
$x_{M0} \times \tilde{y}_M$				−0.172 (0.249)		−0.267 (0.266)
$x_{J0} \times \tilde{y}_I$				0.084 (0.233)		0.026 (0.239)
$x_{C0} \times \tau_{abstract}^{AD}$					0.466 (0.158)	0.264 (0.173)
$x_{M0} \times \tau_{manual}^{AD}$					0.419 (0.282)	0.730 (0.285)
$x_{J0} \times \tau_{routine}^{AD}$					−0.199 (0.162)	−0.130 (0.163)
Tenure	0.238 (0.025)	0.239 (0.026)	0.234 (0.025)	0.235 (0.025)	0.240 (0.026)	0.233 (0.025)
Experience	0.268 (0.014)	0.299 (0.014)	0.266 (0.014)	0.269 (0.014)	0.300 (0.014)	0.267 (0.014)
Years of education	0.270 (0.082)	0.336 (0.084)	0.274 (0.081)	0.271 (0.081)	0.329 (0.083)	0.266 (0.080)
Constant	4.436 (0.094)	4.524 (0.095)	4.448 (0.094)	4.552 (0.145)	4.600 (0.110)	4.562 (0.149)
Observations	224,417	224,417	224,417	224,417	224,417	224,417
Adjusted R^2	0.373	0.339	0.375	0.376	0.340	0.378

Notes: Standard errors clustered at the individual level. τ^{AD} are the task measures from Autor and Dorn (2013b).

any structure on similar occupations (see, for example, Kambourov and Manovskii 2009a, b). This is very much in the spirit of a generalized Roy model. This would be more flexible from a purely empirical vantage point, but substantially more demanding to model, as it would imply that the worker's state space has at least the dimension of the number of occupations. Additionally, it would require the identification and estimation of the latent stochastic processes for each occupation specific skill, and the correlation structure. Authors who have pursued this modeling strategy have typically restricted attention to a small number (3 to 5) of very coarse

occupation groups such as, white, blue and pink collar, service, professional, etc., and assumed independence of ability across occupations (see, for example, Keane and Wolpin 1997; Lee and Wolpin 2006; Sullivan 2010).

In Table 2 we consider the empirical content of our skill measures relative to occupation fixed effects, again in terms of the descriptive wage regression. Column 1 regresses log wages on our vector of skill measures for workers, the skill requirements of their occupation, and the interactions. These coefficients are all used as moments in our estimation.²³ In column 2 we drop our occupation skill demand measures and replace them with occupation fixed effects at the one-digit level.²⁴ In column 3 we include the occupation fixed effects and the interactions between our worker skill measures and our occupation skill requirement measures. There are several things to note. First, the adjusted R^2 is higher when we use our measures than when we use the one-digit occupation fixed effects. Second, adding both the interaction of worker skills with occupation skill requirements and occupation fixed effect improves the fit substantially over the fixed effects regression and only marginally over the skill requirement regression. Additionally, the coefficients on the interaction terms are all significantly different from zero, even when occupation fixed effects are included.

In terms of explaining wage variation, our representation with three worker skills and three skill requirements provides a better fit than including one-digit occupation fixed effects. One-digit occupations are arguably too coarse. Columns 4 and 5 replicate the regressions in columns 2 and 3 with three-digit occupation fixed effects. In this case the adjusted R^2 from the fixed effects regression is higher than using only our skill measures, but the coefficient on the interaction of workers cognitive skills and the cognitive skill requirements of an occupation are still highly statistically significant. There is clearly additional content in our measure of skills and skills requirements. That said, a Roy model of occupation choice at the three digit level would require over 700 additional state variables, each with a latent stochastic process to be estimated. This is well beyond what we consider feasible. An alternative would be to consider a Lucas and Prescott (1974) type islands model (see Kambourov and Manovskii 2009a). This has some attraction, but comes at the cost of giving up any notion of workers being more likely to move to similar occupations.

In columns 6–10 we repeat this exercise including worker fixed effects. Our structural model has unobserved worker heterogeneity that affects the wage multiplicatively, so the fixed effect regression is coherent here too. The main takeaway from these regressions is that even after we include worker fixed effects and three-digit occupation fixed effects, the coefficients on the interaction of worker skills and occupation skill requirements are statistically significant, with the exception of interpersonal skills (we will return to this point when interpreting the structural estimates in Section VB).

²³The coefficient estimates differ slightly from column 4 of Table 1 due to the fact that we lose some observations in Table 1 when matching our occupations to the AD task measures.

²⁴These are the 1970 Census codes used by the NLSY79, corresponding to the 12 categories: Professional, Technical, and Kindred Workers; Managers and Administrators, except Farm; Sales Workers; Clerical and Unskilled Workers; Craftsmen and Kindred Workers; Operatives, except Transport; Transport Equipment Operatives; Laborers, except Farm; Farmers and Farm Managers; Farm Laborers and Farm Foremen; Service Workers, except Private Household; Private Household Workers.

TABLE 2—OCCUPATION AND INDIVIDUAL FIXED EFFECTS

log wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
x_{C0}	−0.036 (0.153)	0.567 (0.116)	−0.130 (0.127)	0.449 (0.105)	−0.144 (0.121)					
x_{M0}	0.014 (0.169)	−0.153 (0.090)	−0.150 (0.107)	−0.124 (0.082)	−0.065 (0.110)					
x_{I0}	0.232 (0.101)	0.311 (0.055)	0.033 (0.067)	0.276 (0.049)	0.105 (0.069)					
\tilde{y}_C	0.041 (0.164)					−0.532 (0.154)				
\tilde{y}_M	0.365 (0.171)					0.561 (0.154)				
\tilde{y}_I	0.395 (0.143)					0.388 (0.148)				
$x_{C0} \times \tilde{y}_C$	0.921 (0.221)		1.161 (0.102)		1.114 (0.123)	1.356 (0.228)		0.731 (0.117)		0.752 (0.116)
$x_{M0} \times \tilde{y}_M$	−0.109 (0.254)		0.202 (0.091)		0.076 (0.110)	−0.279 (0.237)		0.279 (0.085)		0.170 (0.088)
$x_{I0} \times \tilde{y}_I$	0.095 (0.233)		0.556 (0.101)		0.350 (0.124)	−0.144 (0.257)		0.304 (0.109)		0.183 (0.112)
Tenure	0.234 (0.025)	0.261 (0.025)	0.242 (0.024)	0.232 (0.023)	0.232 (0.023)	0.142 (0.019)	0.121 (0.019)	0.138 (0.019)	0.115 (0.018)	0.134 (0.018)
Experience	0.269 (0.014)	0.289 (0.014)	0.264 (0.013)	0.257 (0.013)	0.244 (0.013)	0.335 (0.013)	0.363 (0.013)	0.334 (0.013)	0.343 (0.013)	0.322 (0.013)
Years of education	0.256 (0.081)	0.321 (0.085)	0.294 (0.080)	0.306 (0.075)	0.289 (0.073)					
Constant	4.603 (0.148)	4.237 (0.194)	4.440 (0.200)	4.579 (0.221)	4.751 (0.248)	5.297 (0.058)	5.303 (0.173)	4.991 (0.185)	5.548 (0.130)	5.332 (0.151)
Occupation fixed effects 1 digit		✓	✓				✓	✓		
Occupation fixed effects 3 digit				✓	✓				✓	✓
Worker fixed effects						✓	✓	✓	✓	✓
Observations	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303
Adjusted R^2	0.374	0.347	0.388	0.430	0.448	0.682	0.677	0.684	0.697	0.701

Note: Standard errors clustered at the individual level.

Skill Gaps and Worker Mobility.—In Table 3 we present reduced-form evidence on how skill gaps affect mobility between a worker's first and second jobs. What we mean by "skill gap" is the gap between measured worker skills and measured job requirements in a worker's first job. We examine how the first-job skill gap, interacted with the amount of time the worker spends in that first job, affects the type of job the worker moves to next. Columns 1–3 present results from a set of regressions where the dependent variable is the cognitive, manual, or interpersonal skill requirement at a worker's second job (denoted \tilde{y}_k^+). The right-hand-side variables are the same in each regression and consist of the worker's initial skill bundle; the squared difference between worker's initial skills and the skill requirements in the worker's first job, interacted with an indicator for whether the worker is under- or over-qualified (i.e., the skill gaps in all three skill dimensions); the duration of the first job spell; and the duration of the first job spell interacted with each skill gap. We exclude from this sample workers who have a period of unemployment between the measurement of \mathbf{x}_0 and their first job.

TABLE 3—EFFECT OF QUALITY AND DURATION OF FIRST JOB ON QUALITY OF SECOND JOB

	\tilde{y}_C^+ (1)	\tilde{y}_M^+ (2)	\tilde{y}_I^+ (3)	\tilde{y}_C^+ (4)	\tilde{y}_M^+ (5)	\tilde{y}_I^+ (6)
x_{C0}	0.650 (0.062)	−0.300 (0.074)	0.460 (0.061)	0.653 (0.062)	−0.299 (0.074)	0.462 (0.061)
x_{M0}	−0.117 (0.062)	0.687 (0.074)	−0.409 (0.061)	−0.119 (0.062)	0.686 (0.074)	−0.409 (0.061)
x_{I0}	0.054 (0.065)	0.013 (0.077)	0.395 (0.064)	0.053 (0.064)	0.012 (0.077)	0.394 (0.064)
$\max\{\tilde{y}_C - x_{C0}, 0\}^2$	3.044 (0.694)	0.998 (0.827)	1.102 (0.686)	3.004 (0.693)	0.975 (0.827)	1.077 (0.686)
$\min\{\tilde{y}_C - x_{C0}, 0\}^2$	−0.677 (0.106)	−0.164 (0.126)	−0.096 (0.104)	−0.689 (0.105)	−0.171 (0.126)	−0.103 (0.104)
$\max\{\tilde{y}_M - x_{M0}, 0\}^2$	−0.171 (0.227)	0.682 (0.270)	−0.450 (0.224)	−0.198 (0.226)	0.666 (0.270)	−0.467 (0.224)
$\min\{\tilde{y}_M - x_{M0}, 0\}^2$	0.226 (0.123)	−0.420 (0.146)	0.190 (0.121)	0.239 (0.123)	−0.413 (0.146)	0.198 (0.121)
$\max\{\tilde{y}_I - x_{I0}, 0\}^2$	−0.049 (0.312)	0.011 (0.371)	0.980 (0.308)	−0.077 (0.311)	−0.005 (0.372)	0.963 (0.308)
$\min\{\tilde{y}_I - x_{I0}, 0\}^2$	0.104 (0.109)	0.026 (0.129)	−0.399 (0.107)	0.109 (0.108)	0.029 (0.129)	−0.396 (0.107)
Duration	0.014 (0.005)	−0.001 (0.005)	0.017 (0.005)	0.015 (0.005)	−0.001 (0.005)	0.017 (0.005)
Duration $\times (\tilde{y}_C - x_{C0})$	0.050 (0.020)	−0.038 (0.023)	0.035 (0.019)	0.053 (0.020)	−0.036 (0.024)	0.037 (0.019)
Duration $\times (\tilde{y}_M - x_{M0})$	−0.003 (0.016)	0.078 (0.019)	−0.025 (0.016)	−0.002 (0.016)	0.079 (0.019)	−0.024 (0.016)
Duration $\times (\tilde{y}_I - x_{I0})$	0.002 (0.013)	−0.001 (0.015)	0.031 (0.012)	0.003 (0.013)	−0.000 (0.015)	0.032 (0.012)
Constant	0.091 (0.040)	0.327 (0.047)	0.161 (0.039)	0.139 (0.047)	0.355 (0.056)	0.191 (0.046)
Controls for occupation-specific wage decile				✓	✓	✓
Observations	528	528	528	528	528	528
Adjusted R^2	0.376	0.276	0.497	0.379	0.276	0.498

Note: Standard errors in parentheses.

The coefficients of interest in each regression are on the interactions between the skill gaps in all three skill dimensions interacted with the duration of the first job spell. This tells us how the average skill- k requirement at the second job is affected by the duration of exposure to the same skill requirement at the first job. Take, for example, two workers who have identical initial skill vectors \mathbf{x}_0 , who both have a first job with identical skill demands $\tilde{\mathbf{y}}$. Assume that these workers are under-qualified for their initial job, $\tilde{y}_C - x_{C0} > 0$. The positive coefficient on duration $\times (\tilde{y}_C - x_{C0})$ in column 1 implies that between two identical workers who have identical first jobs for which they are under-qualified on the cognitive skill dimension, the worker who is exposed to the first job for a longer duration will, on average, move to a second job with a higher cognitive skill demand than the shorter duration worker. Similarly, if the workers are over-qualified in the cognitive dimension in their first job, the one who is exposed to this job for a longer duration will, on average, move to a second job with a lower cognitive skill requirement than the short duration worker.

Under- or over-qualification on the manual or interpersonal skills in the first job has no effect on the expected cognitive skill demand in the second job. We see the same pattern for manual and interpersonal skills in columns 2 and 3, with the exception that the cognitive and manual gaps also influence the interpersonal skill requirements of the second job, although with opposite signs.²⁵

Taking Stock.—There are clear patterns in the raw data: the interaction between worker skills and skill requirements affect wages, and how long a worker spends at an initial job where she/he is under- or over-qualified affects the set of jobs she will move to. We next turn to estimation of the structural model to provide a coherent interpretation of these patterns.

IV. Estimation

We estimate the model by indirect inference. To this end, the first step is to simulate a panel that mimics our estimation sample. We first describe the simulation protocol, then discuss the moments we choose to match in the estimation as well as identification of the model.

A. Simulation

Solution Method.—The model has a convenient recursive structure. Equations (2) and (3) can be solved jointly for $U(\mathbf{x})$ and $P(\mathbf{x}, \mathbf{y})$ in a first step. Wages are then obtained from the combination of (4) and the assumption of Bertrand competition: the surplus share $\sigma(\mathbf{x}, \mathbf{y}, \mathbf{y}')$ obtained by a type- \mathbf{x} worker playing off employers \mathbf{y} and \mathbf{y}' (with $P(\mathbf{x}, \mathbf{y}) > P(\mathbf{x}, \mathbf{y}')$) against each other solves (1), and the wages that follow from that renegotiation solve (5). Finally, given those value functions, a cohort of workers can be simulated as we now describe.

Simulation Protocol.—We simulate a cohort of N workers (indexed by $i = 1, \dots, N$) over $T = 300$ months (indexed $t = 0, \dots, T - 1$) using a discrete-time approximation of our model. All workers start out in period $t = 0$ endowed with an initial skill bundle $\mathbf{x}_{i0} = (x_{C,i0}, x_{M,i0}, x_{I,i0}, x_{T,i0})$ drawn from the distribution $\nu(\cdot)$, and in an initial labor market state (unemployed or employed in a job with attributes \mathbf{y}_{i1} under some initial labor contract giving him a share σ_{i1} of the surplus associated with his job) determined as described below. In each subsequent period $t = 1, \dots, T - 1$, we update each worker's skill bundle iteratively using the solution to $\dot{\mathbf{x}}_{is} = \mathbf{g}(\mathbf{x}_{is}, \mathbf{y}_{i,t-1})$ over $s \in [t - 1, t]$ given the initial condition $\mathbf{x}_{i,t-1}$, and where $\mathbf{y}_{i,t-1}$ is the skill requirement vector of the worker's current job (normalized to zero for unemployed workers). We then let any employed worker be randomly hit by a job destruction shock (probability δ) or an outside offer (probability λ_1). Any employed worker hit by a job destruction shock starts the following

²⁵ Groes, Kircher, and Manovskii (2015) finds that the probability of occupational mobility is U-shaped in terms of a worker's occupation-specific wage ranking. In columns 4, 5, and 6 we control for a workers' occupation-specific wage decile (see Appendix Section AB for details). This has essentially no effect on the results for the effect of the quality and duration of a match and mobility.

period as unemployed. Any employed worker receiving an outside offer draws job attributes \mathbf{y}' from the sampling distribution $\Upsilon(\cdot)$ and, depending on the comparison between the value of their current job $P(\mathbf{x}_{it}, \mathbf{y}_{i,t-1})$ and that of their outside offer $P(\mathbf{x}_{it}, \mathbf{y}')$, either accepts the offer (in which case their job attribute vector gets updated to $\mathbf{y}_{it} = \mathbf{y}'$), or stays in their job, with or without a contract renegotiation. In each case, the worker's period- t wage $w_{i,t}$ is updated according to equation (5). Symmetrically, we let any unemployed worker draw a job offer (probability λ_0) with job attributes $\mathbf{y}' \sim \Upsilon(\cdot)$, which the worker accepts if and only if $P(\mathbf{x}_{it}, \mathbf{y}') \geq U(\mathbf{x}_{it})$. Again, the worker's wage is updated.²⁶

To set the initial ($t = 0$) condition, we simulate the model over a “presampling” period, starting from a situation where all workers are unemployed. We then run the simulation as described above, shutting down skill updating and layoffs. We stop the presampling simulation when the simulated non-employment rate reaches a value of 35 percent (the observed non-employment rate on labor market entry in our NLSY sample), and take the current state of the sample at that point as the initial condition. This presampling period is necessary to generate a latent outside option for each worker who is already employed at the time of the survey.

Each simulation produces an $N \times T$ (balanced) panel of worker data with the same format as our estimation sample. The simulated sample keeps track of each worker's employment status, labor market transitions, wages w_{it} , skill bundle \mathbf{x}_{it} , and job attributes \mathbf{y}_{it} .

Model Parameterization.—We use the specification introduced in Subsection IIB which, because it affords closed-form solutions, considerably reduces the computational burden. The skill adjustment, production, and disutility of work functions are specified as in (6), (8), and (9) respectively.

We interpret (x_C, x_M, x_I) and (y_C, y_M, y_I) as the model counterparts of the cognitive and manual skill indices we constructed from our combination of O*NET and NLSY data as explained in the previous section. The joint distribution of initial cognitive, manual, and interpersonal worker skills $(x_C(0), x_M(0), x_I(0))$ is fully observed in the data, and requires no parameterization. General worker efficiency grows along with potential experience t at a constant rate g . In addition, we allow it to be correlated in an unrestricted way with initial cognitive, manual, and interpersonal skills $(x_C(0), x_M(0), x_I(0))$, as well as education:

$$(11) \quad x_T(t) = \exp(g \cdot t + \zeta_S \cdot \text{YEARS_OF_EDUCATION} \\ + \zeta_C x_C(0) + \zeta_M x_M(0) + \zeta_I x_I(0) + \varepsilon_0),$$

where the ζ terms are coefficients and ε_0 is an uncorrelated unobserved heterogeneity term such that the mean of e^{ε_0} is normalized to 1. Given the model's structure, this makes e^{ε_0} an uncorrelated mixing variable that multiplies all individual wages and values. In particular, observed log-wages $\ln w$ are such that $\ln w \stackrel{d}{=} \ln w|_{\varepsilon_0=0} + \varepsilon_0$, where $\stackrel{d}{=}$ denotes equality in distributions, and $w|_{\varepsilon_0=0}$ denote

²⁶Note that, in the simulation, we shut down sample attrition (which in the model occurs at rate μ). Attrition is random in the model, the only impact would be to reduce the simulated sample size, which we can usefully avoid.

simulated wages under the assumption that all workers have $\varepsilon_0 = 0$. We can thus estimate the model abstracting from this particular heterogeneity (i.e., assuming $\varepsilon_0 = 0$ for all workers), then retrieve the distribution of ε_0 by deconvolution.

Finally, we specify the skill requirements (y_C, y_M, y_I) as simple transforms of the skill requirement indices $(\tilde{y}_C, \tilde{y}_M, \tilde{y}_I)$ constructed from the O*NET data as described in Section III. This is to allow for the possibility that our constructed \tilde{y} s might not be exactly aligned with our measured worker skills \mathbf{x} (see the discussion in Section III). Specifically, we assume that $y_k = \tilde{y}_k^{\xi_k}$, with $\xi_k > 0$, thus ensuring that y_k is an increasing transformation of \tilde{y}_k that stays in the unit interval. We then approximate the joint sampling distributions of job attributes $\Upsilon(\mathbf{y})$ using a Gaussian copula and Beta marginals with skill-specific parameters (η_k^1, η_k^2) , $k = C, M, I$. The rank correlation parameters $(\rho_{CM}, \rho_{CI}, \rho_{MI})$ of the Gaussian copula are to be estimated, together with the parameters of the three marginals.

B. Targeted Moments

The specification of our model laid out in Subsection IVA involves the parameter vector described earlier in this paper and summarized in Table A1 in the Appendix. Among those parameters, we fix the discount rate r and the sample attrition rate μ to “standard” values (the monthly equivalent of 10 percent per annum for r , and 0.002 for μ , implying an average working life of 42 years). As explained before, the joint distribution of initial cognitive, manual, and interpersonal worker skills $(x_C(0), x_M(0), x_I(0))$ is observed in the initial cross section of our estimation panel. Finally, the job destruction rate δ has a direct empirical counterpart, namely the sample average job loss (E2U) rate.²⁷ With this subset of parameters estimated, or calibrated, in a preliminary step, we are left with a 39-dimensional parameter vector to estimate (summarized in Table A1 in the Appendix). We estimate these parameters by matching the following set of moments: (i) sample mean U2E rate, (ii) mean E2E rate profile (summarized by average E2E rates over six consecutive equal-length subsets of the observation window), (iii) mean and standard deviation of the marginal cross-sectional distributions of current job attributes $\tilde{\mathbf{y}}_{it}$ among employed workers at a selection of sampling dates,²⁸ (iv) pairwise correlations of skill requirements, $\text{corr}(\tilde{y}_{k,it}, \tilde{y}_{k',it})$, $k' \neq k$, $(k, k') \in \{C, M, I\}^2$ among jobs held by employed workers at a selection of sampling dates, (v) correlations of initial worker cognitive, manual, and interpersonal skills and the skill requirements of jobs held, $\text{corr}(x_{k,i0}, \tilde{y}_{k,it})$, $k = C, M, I$ at a selection of sampling dates, (vi) coefficients of a regression of log wages $\ln w_{it}$ on initial skills \mathbf{x}_{i0} , current job attributes $\tilde{\mathbf{y}}_{it}$, interactions $x_{k,i0} \times y_{k,it}$ for $k = C, M, I$, tenure, experience, and years of education (i.e., the regression in column 1 of Table 2). We drop the first simulated wage out of unemployment for the wage regression.²⁹ The model-based moments are computed

²⁷ Figure 5 suggests that the job loss rate is not exactly constant over a worker’s life cycle. We abstract from this feature of the data.

²⁸ In practice, we compute those moments at six dates corresponding to 2.5, 5, 7.5, 10, 12.5, and 15 years into the sample.

²⁹ In this version of the sequential auction model, in which workers are risk-neutral and have no bargaining power, workers tend to accept very low wages upon exiting unemployment, to “buy their way” onto the job ladder. As soon as a worker receives her/his first outside offer the wage will jump. We drop the initial wage out of unem-

from simulated samples of $N = 18,400$ workers, ten replicas of the initial NLSY cross section.

C. Identification

Appendix AE formally discusses identification of the model laid out in Section II (given the parameterization described in Subsection IIB) from a dataset with the structure and contents described in Section III. In this subsection we summarize the main sources of information that identify the various components of our model.

The levels of wages conditional on education, experience, initial skills, and (observed) job skill requirements identify the returns to education and initial skills (the parameters ζ_k , $k = C, M, I$), the wage trend (parameter g), and the baseline returns to job skill requirements (parameters α_k and α_{kk}). The (production/utility) costs of mismatch and the speed of human capital accumulation or decay (parameters $\kappa_k^{u/o}$ and $\gamma_k^{u/o}$) are identified from comparisons of the sets of job types \mathbf{y} that are acceptable to workers with equal initial skills $\mathbf{x}(0)$, but have experienced different employment histories. Knowledge of any worker's initial skill bundle $\mathbf{x}(0)$ and full labor market spell history, combined with the knowledge of the skill adjustment process (parameters $\gamma_k^{u/o}$) then enables us to construct the full path of skill bundles $\mathbf{x}(t)$ for all workers in the sample. The set of job offers accepted by unemployed workers with given skill bundle \mathbf{x} then identifies the sampling distribution $\Upsilon(\mathbf{y})$ over the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) \geq U\}$, so that $\Upsilon(\mathbf{y})$ is identified over the union of all such sets for all skill bundles \mathbf{x} observed in the sample (that is, $\Upsilon(\mathbf{y})$ is identified at all skill requirement levels \mathbf{y} that are acceptable by at least some worker types). Finally, the offer arrival rates λ_0 and λ_1 are identified, conditionally on the rest of the model, from sample U2E and E2E transition probabilities.

Although the exact arguments used in Appendix Section AE to establish identification are not literally taken up in the practical estimation protocol, the information contained in the moments we use for estimation (listed in Subsection IVB) does echo those arguments. In particular, the cross-section wage regression coefficients that we seek to replicate contain the information needed to identify the parameters α_k , α_{kk} , ζ_k , and g . Moreover, the various moments of the joint distribution of initial worker skills and current job skill requirements convey information about the set of matches that are acceptable to a given worker type, which is used to identify $\kappa_k^{u/o}$, $\gamma_k^{u/o}$, and ultimately the sampling distribution $\Upsilon(\cdot)$.

The descriptive wage regression in Table 1 provides some reduced-form evidence of which skills requirements are most important in production, and illustrates where some of the identification of the production function is coming from. In column 1, log wage is regressed on our measures of initial worker skill bundles and the skill requirements of the job they are currently in. When entered linearly, both cognitive and interpersonal skills of the workers affect the wage positively, as do the cognitive, manual, and interpersonal skill requirements of the job. A worker's initial level of manual skills does not have a statistically significant effect on wages. In column 4 we add to this regression interaction terms between a worker's measured skill

ployment so as not to bias our estimate of human capital accumulation due to the large wage change at the very beginning of an employment spell. We return to this issue in online Appendix Section B.B1.

and the skill requirement of the job ($x_{k0} \times \tilde{y}_k$), for $k \in \{C, M, I\}$. An informative pattern emerges. The coefficient on the interaction of a worker's initial cognitive skills and the cognitive skill requirement of the job is strongly positive, while the coefficients on x_{C0} and \tilde{y}_C entered separately become small and statistically insignificant. Interestingly, the coefficient on the interaction of worker skills and job skill requirements for manual and for interpersonal dimensions are small and statistically insignificant, and the coefficient on these skills and skill requirements entered individually remain essentially unchanged from column 1. In terms of identification of our model, these moments will imply a production function in which employing a worker who meets the cognitive skill requirements for a job is very important, much more so than meeting the manual or interpersonal skill requirements.

Identification of the human capital accumulation process for each skill type comes from how past jobs affect future opportunities. We presented some direct evidence on this in Table 3. Workers who are employed at a job in which they are initially under-qualified in a particular skill dimension move, on average, to jobs that require even higher levels of this skill requirement, and this is increasing in duration of the first match. We see clearly that the initial gap between a worker's skill and the job skill requirements, interacted with duration of the first job affects the set of jobs the worker accepts in the future. This dependence of accepted jobs in the future on early career skill gaps is informative on the extent to which workers learn different skills depending on the skill requirements of the job they are in, and how long they stay at the job.

In practice, our chosen moments ensure precise "local" identification of the model's parameters, in the sense that the distance between data-based model-predicted moments has a clear local minimum at the estimated parameter value.³⁰

Several caveats on identification are worth mentioning here. First, our model implies that unobserved heterogeneity ε_0 affects wages, but does not affect a worker's decision to accept or reject a job offer. To the extent that workers do select jobs based on unobserved characteristics, our model will tend to overestimate the extent of mismatch and overstate the implied role of search frictions. Second, identification is parametric. For example, it requires that the shape of the production function estimated at the boundary of the acceptance set is informative about the shape of the production function in the interior.

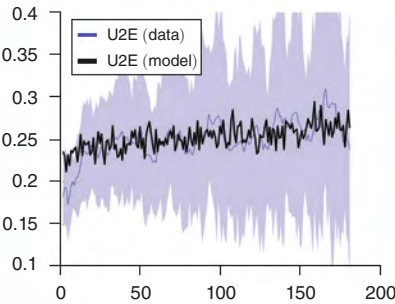
V. Results

A. Model Fit

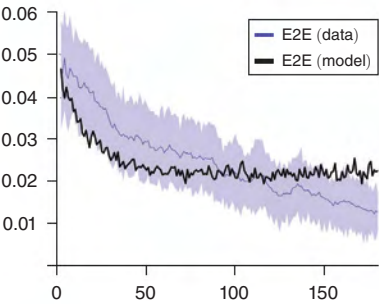
Figure 2 illustrates various aspects of the fit. All time series in Figure 2 are plotted over a period of 15 years (180 months, i.e., the sample window used for estimation), together with 95 percent confidence bands (based on 1,500 bootstrap replications) around the data series. Panel H further shows the fit in terms of the descriptive wage regression discussed in Section III.

³⁰ We present additional measures of the local sensitivity of the parameter estimates to the data moment in online Appendix Section B.B2.

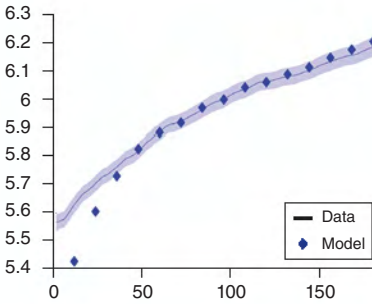
Panel A. U2E rate



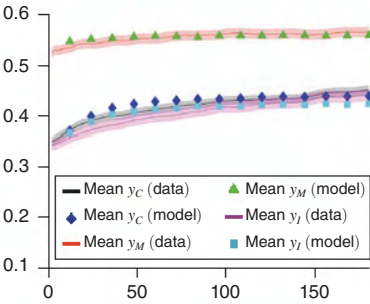
Panel B. E2E rate



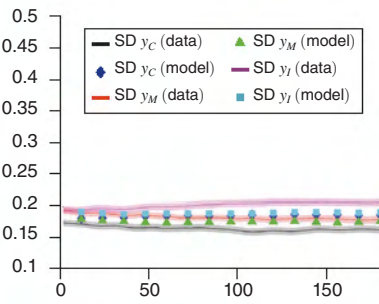
Panel C. log wage/experience profile



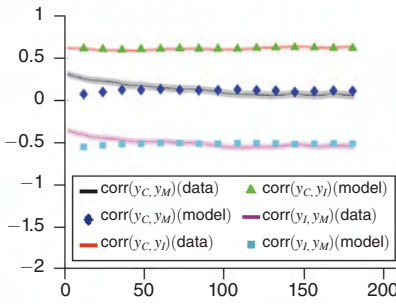
Panel D. Cross-sectional mean job attributes



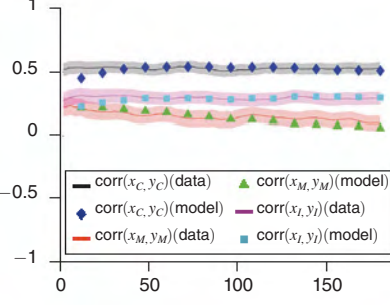
Panel E. Cross-sectional standard deviation of job attributes



Panel F. Correlation of job attributes



Panel G. Correlation of job and worker attributes



Panel H. Descriptive (log) wage regression

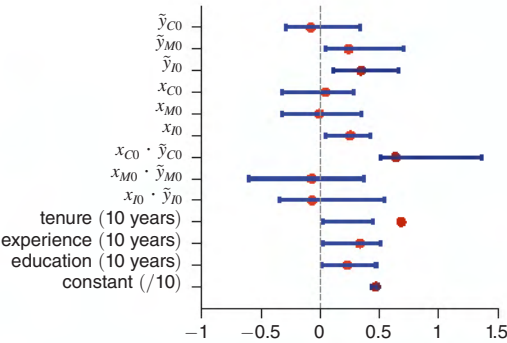


FIGURE 2. MODEL FIT

The model fits both the non-employment exit rate (U2E, panel A) and the job-to-job transition rate (E2E, panel B) reasonably well, considering the restriction to constant job contact rates λ_0 and λ_1 over the life cycle. The decline of E2E rates with experience is correctly captured by the model (it occurs as a consequence of workers gradually settling into jobs to which their skills are better suited, both because they sort into better matches over time and because their skills adjust to whatever job they are in at any given time), even though the model overstates both the initial speed of that decline and the level of the E2E rate at high levels of experience. All of the discrepancies between data and model in panels A and B of Figure 2 are largely due to our restriction to experience-invariant contact rates, λ_0 and λ_1 .

The sample average wage/experience profile is shown in panel C of Figure 2. We do not directly target that particular profile in the estimation, yet the model captures it reasonably well, despite a tendency to overstate its concavity.

Panels D through G show the time-profiles of various fitted cross-sectional moments of the joint distribution of workers' initial skills ($x_{C,i0}, x_{M,i0}, x_{I,i0}$) and current job attributes ($\tilde{y}_{C,it}, \tilde{y}_{M,it}, \tilde{y}_{I,it}$) in the population of employed workers, at a selection of experience levels. The model gets the levels and experience profiles of those cross-sectional moments roughly right. In particular, the model captures the rise in average cognitive and interpersonal job attributes, although it tends to overstate the concavity of the experience profiles of those two job attributes at low levels of experience. It also replicates the near constancy of average manual job attributes. The model further captures the experience profiles of correlations between the various job attributes (panel F), and between job requirements in each skill category and the corresponding initial worker skill (panel G), again despite slight discrepancies between data and simulation at low levels of experience. Finally, while the model correctly predicts a flat experience profile for the standard deviations of all three job attributes, it does a little less well fitting the levels of those standard deviations, overstating the standard deviations of cognitive skill requirements and understating that of interpersonal skill requirements by about 10 percent at all levels of experience.

We finally turn to the model's ability to replicate the pooled cross-section wage regression shown in column 1 of Table 2. Panel H of Figure 2 shows, for each coefficient of the regression, the bootstrap 95 percent confidence interval of the empirical estimate, together with the structural model-based point estimate (the large dots). The structural estimates are all well within the empirical confidence bounds, with the exception of the returns to job tenure which the model has a tendency to overstate.

B. Parameter Estimates

Table 4 shows point estimates of the model parameters with asymptotic standard errors in parentheses below each estimate.³¹ There is little to say about the offer arrival and job destruction rates, which are within the range of standard estimates

³¹The covariance matrix of the parameter vector Θ is estimated as $(G^T G)^{-1} G^T \Omega (G^T G)^{-1}$, where $G = E[\partial \mathbf{m} / \partial \Theta^T]$, the (expectation of the) Jacobian matrix of the moment function $\mathbf{m}(\Theta)$, is obtained by numeri-

TABLE 4—PARAMETER ESTIMATES

Production function						
α_T 137.5 (17.0)	α_C 140.3 (24.8)	α_M 64.4 (10.1)	α_I 92.4 (11.1)	α_{CC} 195.6 (24.2)	α_{MM} 10.7 (2.40)	α_{II} 15.4 (4.29)
Production function ^a			Disutility of work ^a			
κ_C^u 5,165.3 (533.5) 45.7	κ_M^u 984.5 (265.4) 3.9	κ_I^u 337.6 (97.4) 4.1	κ_C^o 54.1 (7.14) 20.9	κ_M^o 409.6 (71.9) 8.8	κ_I^o 171.9 (23.9) 5.1	
Skill accumulation function ^b						
γ_C^u 7.7e−3 (8.4e−4) 7.54	γ_C^o 2.1e−3 (5.3e−4) 27.3	γ_M^u 3.4e−2 (3.1e−3) 1.70	γ_M^o 7.7e−3 (6.4e−4) 7.51	γ_I^u 1.0e−3 (3.8e−3) 55.8	γ_I^o 5.8e−7 (1.4e−3) 99,407	g 2.3e−3 (5.0e−4)
General efficiency				Transition rates ^c		
ζ_S 2.4e−2 (0.031)	ζ_C 0.18 (0.501)	ζ_M −0.17 (0.521)	ζ_I 0.20 (0.261)	λ_0 0.39 (0.011)	λ_1 0.16 (1.2e−3)	δ 2.1e−2 (3.3e−7)
Sampling distribution ^d						
ξ_C 1.21 (0.038)	ξ_M 0.79 (0.038)	ξ_I 0.88 (0.040)	ρ_{CM} 0.14 (0.019) 0.13	ρ_{CI} 0.73 (0.011) 0.71	ρ_{IM} −0.44 (0.019) −0.42	
Sampling distribution ^e						un. inc.
η_C^1 1.22 (0.066)	η_C^2 2.86 (0.095)	η_M^1 2.15 (0.143)	η_M^2 2.76 (0.117)	η_I^1 0.93 (0.085)	η_I^2 2.96 (0.124)	b 137.5 (17.0)
	0.20 0.30		0.44 0.20		0.24 0.19	

Notes:

^a Percent surplus loss caused by deviating from output-maximizing match by 1 SD of Υ at mean \mathbf{x} in italics.^b Half-life in years in italics.^c δ estimated in first step.^d Implied correlations in italics.^e Means and standard deviations in italics.

on US data, even though the ratio $\lambda_1/\lambda_0 \simeq 0.42$ is on the high end of that range. Overall job productivity is increasing in all cognitive, manual, and interpersonal skill requirements, with the loading on cognitive skills between 1.5 and 2 times as large as the ones on manual and interpersonal skills. Job skill requirements are complementary to the corresponding worker skills (α_{CC} , α_{MM} , and α_{II} are all positive), although complementarity is an order of magnitude stronger in the cognitive than in the other two skill dimensions.

Overall worker efficiency x_T is positively associated with a high initial endowment in cognitive and interpersonal skills ($\zeta_C, \zeta_I > 0$), while initial manual skills are

cal differentiation and where Ω , the covariance matrix of the moment function, is estimated by resampling the data 1,500 times.

negatively correlated with x_T ($\zeta_M < 0$). One additional year of education increases efficiency by 2.4 percent (ζ_S). However, one should bear in mind that education is positively correlated with initial cognitive and interpersonal skills and (weakly) negatively correlated with initial manual skills in the sample. The value of ζ_S taken in isolation therefore understates the overall returns to education. Also note that the parameters ($\zeta_S, \zeta_C, \zeta_M, \zeta_I$) are imprecisely estimated.

The employment of an under-qualified worker in any skill dimension is costly in terms of output, yet the output loss caused by this type of mismatch is by far most severe in the cognitive dimension and least severe in the interpersonal dimension. The utility cost of being under-matched, i.e., the surplus cost of the worker being *over*-qualified, is positive in all dimensions, but generally much smaller than the corresponding surplus (production) cost of *under*-qualification. To give a sense of the orders of magnitude involved, the numbers in italics below the estimates of the various κ 's give the percentage flow-surplus cost of deviating by one standard deviation of the sampling distribution Υ from the output-maximizing match for a worker with the mean skill level in the distribution of initial skills \mathbf{x}_0 .³²

The correlation patterns between skill requirements in the sampling distribution ($\rho_{CM} > 0, \rho_{CI} > 0, \rho_{IM} < 0$) suggests that jobs requiring high levels of cognitive skills also tend to require high levels of skills in at least one of the other two dimensions, particularly interpersonal. Jobs with a high manual content, however, tend to have low interpersonal requirements. The next section offers a more complete view of the sampling distribution.

All three types of skills are accumulated faster (by an under-qualified worker) than they are lost (by an over-qualified worker), i.e., $\gamma_k^u > \gamma_k^o$ for $k = C, M, I$. Yet apart from that common property, patterns of skill adjustment differ vastly between skill dimensions. Manual skills adjust much faster than cognitive skills. Cognitive skills are very persistent (i.e., not easily accumulated or lost) with a half-life of 7.5 years to learn and 27.3 years to forget. The half-life of manual skills is much shorter, about 20 months to acquire and 7.5 years to lose. Interpersonal skills essentially do not adjust over a worker's typical horizon and can, to a good approximation, be treated as fixed worker traits, determined prior to labor market entry.

Perhaps the clearest message from those estimates is that the model sees cognitive, manual, and interpersonal skills as very different productive attributes. Manual skills have relatively low returns and adjust quickly in both directions, cognitive skills have much higher returns, but are much slower to adjust, especially upward. Interpersonal skills have lower returns than cognitive skills, but higher than manual skills (including through their effect on overall worker efficiency), and are essentially fixed over a worker's lifetime. Finally, skill mismatch is most costly in the cognitive dimension and in the "underskilled" direction (i.e., when the worker has lower skills than the job requires).

³² Denoting the mean of N as $\mathbf{x}^m = (x_C^m, x_M^m, x_I^m)$, the output-maximizing match is with $\mathbf{y}^* = \arg\max_{\mathbf{y}} f(\mathbf{x}^m, \mathbf{y})$. Then, denoting the standard deviations of the marginals of Υ as $(\sigma_C, \sigma_M, \sigma_I)$, the percentage reported in italics below the estimate of κ_C^u in Table 4 is $100 \times \left[1 - \left(P(\mathbf{x}^m, \mathbf{y}^* + (\sigma_C, 0, 0)) - U(\mathbf{x}^m) \right) / \left(P(\mathbf{x}^m, \mathbf{y}^*) - U(\mathbf{x}^m) \right) \right]$.

C. Skill Mismatch, Skill Changes, and Sorting

Distributions of Skills and Skill Requirements.—The broad question of skill mismatch can be understood in many different ways. One aspect of that question is the alignment (or lack thereof) between the skills that workers are equipped with when they leave education, the distribution $N(\cdot)$ of initial worker skill bundles $\mathbf{x}(0)$, in the parlance of the model, and the firms' skill requirements, the model counterpart of which is the sampling distribution $\Upsilon(\mathbf{y})$.

The top row of Figure 3 (panels A, B, and C) show the marginal sampling distributions of pairs of job attributes, integrating out one skill dimension at a time.³³ The second (panels D, E, and F) do the same for the distribution of initial skills among labor market entrants, $N(\cdot)$. Plots of the sampling distribution suggest that labor demand is concentrated around jobs with intermediate to high manual skill requirements (y_M around 0.5 to 0.6), and modest levels of cognitive and interpersonal skill requirements (y_C and y_I around 0.2). A visual comparison of the top two rows of Figure 3 further suggests that labor market entrants are, on average, endowed with levels of manual skills that roughly coincide with what the sampling distribution suggests employers are looking for, but also seem to have much higher levels of cognitive and interpersonal skills than is required in most jobs.

Of course, there are narrow limits to the amount of information conveyed by a visual comparison of the sampling and initial skill distributions. First, discrepancies between those two distributions are not entirely surprising: the population of workers whose skills are represented in panels D–F of Figure 3 is a cohort of relatively young workers. As such, their skills may not be representative of those in the entire active workforce. By contrast, at least under random search, the sampling distribution addresses all workers, the majority of which are from older cohorts among which the skill distribution may be quite different. Second, work skills and job attributes are combined into the surplus function (10), and what relative shapes of Υ and N constitute a “good fit” in the sense of surplus maximization is not visually obvious. For example, as discussed above, parameter estimates show that being over-qualified in the manual dimension is costly relative to being over-qualified in any of the other two skill dimensions, while being under-qualified in the cognitive dimension is very costly. The apparent excess of cognitive skills in the population of labor market entrants may be a sensible response to those features of the technology. We consider specific surplus-based measures of mismatch in the next section. Third, matching occurs gradually over time, with workers sampling repeatedly from Υ , while their skills evolve over time.

We now investigate this last point. Panels G–L of Figure 3 show how the distribution of worker skills in the model changes as the cohort of workers accumulates experience. The evolution is clearly toward workers gaining cognitive skills and maintaining their manual skills on average (while, as we saw earlier, interpersonal skills hardly adjust at all over a worker's lifetime). This can be explained by the fact that jobs with high cognitive skill requirements are intrinsically more productive (the estimated weights on y_C and $x_C \cdot y_C$ in the production function, α_C and α_{CC} , are an

³³For example, Figure 3, panel A is a plot of $\int v(y_C, y_M, y_I) dy_I$ against (y_C, y_M) .

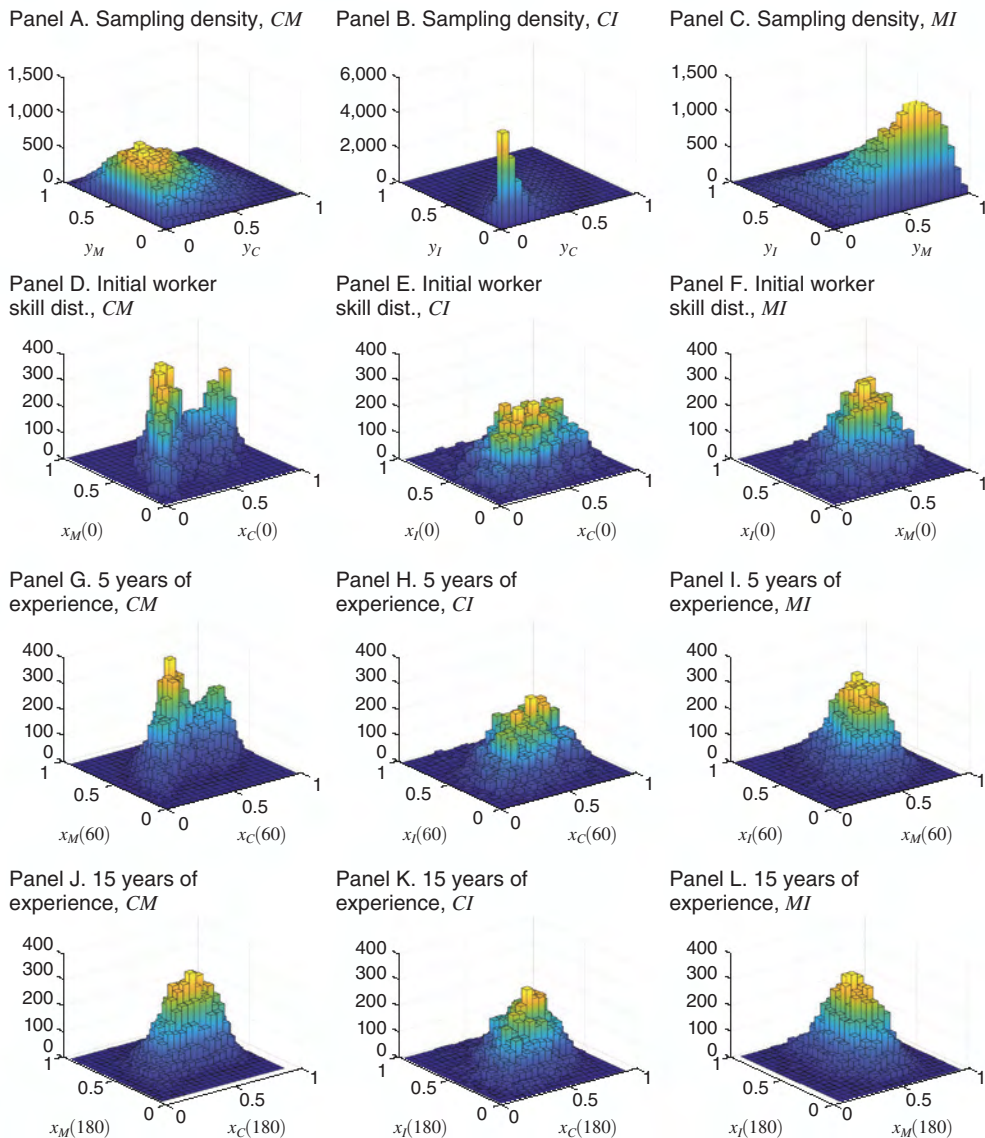


FIGURE 3. DISTRIBUTION OF SKILL REQUIREMENTS AND EVOLUTION OF WORKER SKILLS WITH EXPERIENCE

order of magnitude larger than their manual and interpersonal counterparts), inducing workers to accept jobs with the highest cognitive content compatible with their level of cognitive skills, while at the same time avoiding costly over-qualification in manual skills. By following this strategy workers tend to maintain or gradually acquire cognitive skills. A second striking feature of Figures 3 is that the already limited degree of specialization apparent in the initial skill distribution (panel A) regresses even further as workers gain experience: the skill distribution becomes increasingly unimodal.

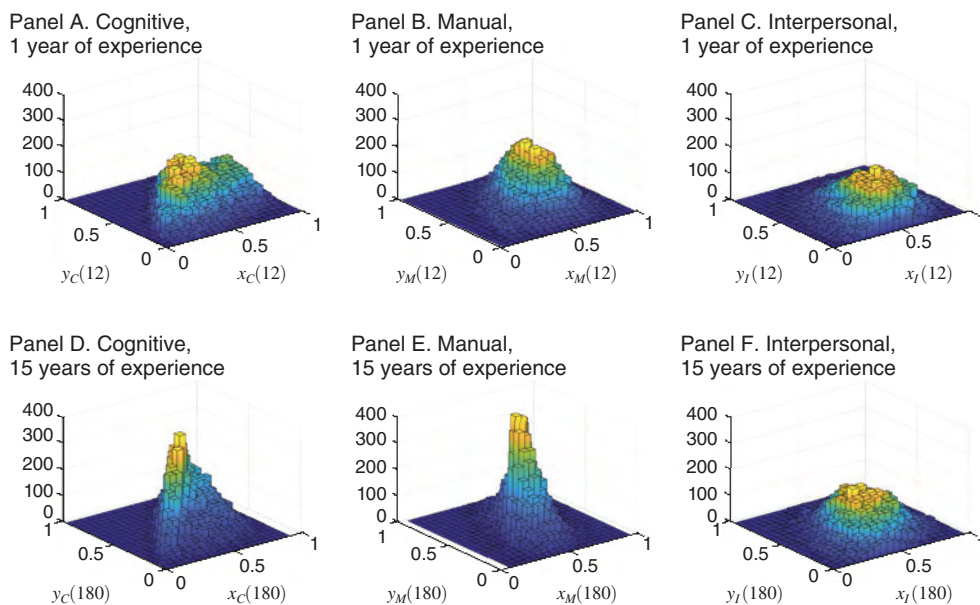


FIGURE 4. SORTING

Skill Sorting and Mismatch.—We next examine the joint distribution of worker skill bundles and job skill requirements among ongoing matches. Figure 4 shows two examples of those joint distributions, among workers who are 1 year into their careers (panels A, B, and C), and among workers with 15 years of experience (panels D, E, and F). Simply eyeballing these histograms gives a distinct impression of positive sorting in all skill dimensions, even at early stages of the working life. Moreover, the “strength” of this positive sorting, as measured by the (inverse of the) conditional dispersion in worker skills for a given level of skill requirement, clearly increases as workers accumulate experience. This results from the combination of workers gradually sorting themselves into jobs for which their skills are better suited, and adjusting their initial skills to their job’s requirements: as can be seen from Figure 4, sorting at 15 years of experience is strongest in the manual dimension (as manual skills adjust quickly), and weakest in the interpersonal dimension (as interpersonal skills do not adjust).

A final feature of Figure 4 is that, while there is largely positive sorting in all skill dimensions, a substantial mass of workers appear “under-matched” in the cognitive dimension, in the sense that their job’s cognitive skill requirement is lower than their own cognitive skill level. By contrast, very few workers are “over-matched” in cognitive skills (and those who are so by a small margin). The trade-off from the perspective of a worker contemplating a job is between the job’s overall productivity (the $\alpha_C y_C$ and $\alpha_{CC} x_C \cdot y_C$ terms in the production function), and any cost of being mismatched. In the case of cognitive skills, the cost of being “over-matched” (or “under-skilled”), measured by κ_C^u , is prohibitively high, even accounting for the fact that output increases much more steeply with y_C than with y_M or y_I (Table 4).

D. Worker Bargaining Power

Our baseline model relies on the sequential auction rent-sharing protocol of Postel-Vinay and Robin (2002), in which the only source of worker rent is Bertrand competition between employers. Upon receiving a job offer, a worker compares the total private values of his/her current labor market state (employment in a given type- y job or unemployment) with that of accepting the job offer (employment in an alternative type- y' job), and never extracts more than the minimum of those two values. This assumption has the implication that workers tend to accept very low wages upon exiting unemployment, to “buy their way” onto the job ladder, after which wages tend to increase very steeply as soon as they receive their first outside offer on their new job. The model thus tends to overstate returns to tenure, as well as returns to experience early on in a worker’s career. The sequential auction protocol further implies that private firm-worker pairs fail to internalize the extra surplus that is created when the worker moves to a competing job for which her/his skills are better suited, as that extra surplus is entirely captured by a third party, namely the worker’s future employer. For the same reason, unemployed workers fail to internalize the surplus associated with them finding a job. Both of those features of the sequential auction model can be mitigated by extending the model to give workers some additional bargaining power that enables them to capture a share $\beta \in [0, 1]$ of the match rent over and above what they receive from sheer Bertrand competition between employers (Cahuc, Postel-Vinay, and Robin 2006; Dey and Flinn 2005). In online Appendix Section B.B1 we investigate this extension and show that, while endowing workers with positive bargaining power is theoretically straightforward in the context of our model, it has a very high cost in terms of computational tractability.

VI. The Determinants of Social Output

We now analyze the determinants of the social value of output in our model economy. Specifically, we focus on the expected present discounted sum of future output produced by a worker from experience t onward (the “experience- t expected career output”). Consider a worker i with experience t , who is either unemployed (denoted by $\ell_{it} = 0$) or employed ($\ell_{it} = 1$) in a job with attributes \mathbf{y}_{it} . The worker has education (years of education) ed_i , initial skill bundle \mathbf{x}_{i0} , unobserved ability ε_{0i} and current skills \mathbf{x}_{it} . Their experience- t expected career output is then defined as

$$(12) \quad Q_{it} = E \left[\int_t^{+\infty} (\ell_{is} [f(\mathbf{x}_{is}, \mathbf{y}_{is}) - c(\mathbf{x}_{is}, \mathbf{y}_{is})] + (1 - \ell_{is}) b(\mathbf{x}_{is})) e^{-(r+\mu)(s-t)} ds \mid \mathbf{x}_{i0}, ed_i, \varepsilon_{0i}, \mathbf{x}_{it}, \ell_{it}, \mathbf{y}_{it} \right].$$

Dispersion in initial skill bundles \mathbf{x}_{i0} , years of education ed_i , and unobserved ability ε_{0i} , along with the random career shocks faced by workers (job mobility, job destruction, and attrition) up to experience t in our cohort of workers induces a cross-section distribution of experience- t career output within the cohort, for any experience level t . We now seek to assess the contributions of those different sources

of randomness to the variance of (the natural logarithm of) Q_{it} for a fixed experience level t .³⁴

To that end, note that, with our specification assumptions, all three functions $f(\cdot)$, $c(\cdot)$, and $b(\cdot)$ are multiplicatively separable in the worker's initial general efficiency, $x_{T,i0}$. This implies that Q_{it} can be decomposed as $\ln Q_{it} = \ln x_{T,i0} + \ln \tilde{Q}_{it}$. Recalling the definition (11) of $x_{T,i0}$, we see that dispersion in $\ln x_{T,i0}$ is induced by dispersion in \mathbf{x}_{i0} , years of education ed_i , and ε_{0i} , while dispersion in $\ln \tilde{Q}_{it}$ is induced by dispersion in \mathbf{x}_{i0} and by random career shocks. Indeed, $\ln \tilde{Q}_{it}$ is independent of years of education and ε_{0i} conditional on \mathbf{x}_{i0} . The variance of career output $\text{var} \ln Q_{it}$ can therefore be decomposed in the following interpretable manner:

$$\begin{aligned} \text{var} \ln Q_{it} &= \text{var} [E(\ln Q_{it} | \mathbf{x}_{i0})] + E[\text{var}(\ln Q_{it} | \mathbf{x}_{i0})] \\ &= \text{var} [E(\ln Q_{it} | \mathbf{x}_{i0})] + E[\text{var}(\ln \tilde{Q}_{it} | \mathbf{x}_{i0})] + E[\text{var}(\ln x_{T,i0} | \mathbf{x}_{i0})] \\ &= \underbrace{\text{var} [E(\ln Q_{it} | \mathbf{x}_{i0})]}_{1: \text{ between } \mathbf{x}_0} + \underbrace{E[\text{var}(\ln \tilde{Q}_{it} | \mathbf{x}_{i0})]}_{2: \text{ shocks}} \\ &\quad + \underbrace{\text{var} \varepsilon_{i0}}_{3: \text{ heterogeneity}} + \underbrace{\zeta_S^2 E[\text{var}(ed_i | \mathbf{x}_{i0})]}_{4: \text{ schooling given } \mathbf{x}_0}, \end{aligned}$$

where the last equality uses (11) to further decompose the conditional variance of $\ln x_{T,i0}$. The first term is the between- \mathbf{x}_0 variance, i.e., the part of the cross-section variance in experience- t career output that is explained by dispersion in initial skill endowments, \mathbf{x}_{i0} . The second term is the part of the cross-section variance in Q_{it} that is attributable to the randomness in a worker's employment history. The third term is the variance attributable to dispersion in unobserved ability ε_{0i} (which, by construction of the model, is uncorrelated to any other source of dispersion). Finally, the fourth term is the residual variance, due to the dispersion in education that is not explained by dispersion in initial skills \mathbf{x}_{i0} . Results from this variance decomposition are gathered in Table 5 for $t = 10$ years of experience. For practical interpretability, we perform our variance decomposition on three subsamples, stratified by education: College and above (16+ years of education, 30.7 percent of the sample), Some college (13–15 years of education, 19.3 percent of the sample), and Non-college (12 years of education or less, 50.0 percent of the sample).

For our entire cohort, almost two-thirds (65 percent) of the variance is accounted for by the vector of initial skill endowments \mathbf{x}_0 , and an additional 19 percent is explained by unobserved heterogeneity ε_0 . The additional information brought about by education given \mathbf{x}_0 is negligible. Taken together, those numbers mean that about 84 percent of the variance in expected career output at 10 years of experience is explained by dispersion in initial conditions. The rest, a little over 16 percent, is due to the randomness associated with job offers and job displacement during the first

³⁴ See Appendix Section AG for details on the computation of Q_{it} .

TABLE 5—DECOMPOSITION OF $\text{var} \ln Q_{it}$

	Percent share of $\text{var} \ln Q_{it}$ due to...			
	Initial skills \mathbf{x}_0 (term 1)	Shocks (term 2)	Heterogeneity ε_0 (term 3)	Education \mathbf{x}_0 (term 4)
Whole sample	65.0	16.4	18.9	0.0
College+	17.2	48.3	35.5	0.0
Some college	27.5	34.2	38.9	0.0
Non-college	37.9	22.4	40.1	0.0

Note: Level of experience: $t = 10$ years.

10 years in the labor market, and the differential skill accumulation that results from this randomness.³⁵ Conditioning on broad levels of education reduces the share of variance explained by initial skill bundles, as those are correlated with education.³⁶ Yet comparison of the bottom three rows of Table 5 suggests that the importance of labor market shocks increases with education, while, vice versa, the importance of initial conditions (the sum of \mathbf{x}_0 and ε_0) decreases with education.

To get a sense of which type(s) of skills generate the most variance in career output, Table 6 reports the shares of $\text{var} \ln Q_{it}$ explained by each separate skill dimension of \mathbf{x}_0 taken in isolation, $\text{var}[E(\ln Q_{it} | x_{0k})]$ for $k \in \{C, M, I\}$. We see that cognitive skills alone can account for almost 60 percent of the variance, much more than what either manual skills (12 percent) or interpersonal skills (19 percent) can account for on their own. Cognitive skills are the most important dimension of initial skills for all education groups, however their relative importance is much smaller for the low-educated: initial cognitive skills explain more than three times as much variance as initial manual skills for the top education group, but only account for 22 percent more variance than initial manual skills for the Non-college group.

Given the dominant role of cognitive skills illustrated in Table 6, a natural question to ask is: what would a model with only a single dimension for skill miss? Table 7 reports the decomposition of $\text{var} \ln Q_{it}$ as obtained from a version of our model with a scalar (one-dimensional) measure of job skill requirements. Specifically, instead of allowing for a three-dimensional job attribute $\mathbf{y} = (y_C, y_M, y_I)$, we only allow for a scalar y , constructed as the first principal component of our initial large set of O*NET measures.³⁷ Symmetrically, we only allow for one dimension of

³⁵ Keane and Wolpin (1997) provides a similar decomposition in which they conclude that 90 percent of the variation is due to initial conditions and 10 percent is due to shocks. The main difference with our approach is that they assume workers are always employed at their best job. From the point of view of our model, this would be equivalent to labeling all of the variation due to mismatch as due to initial skills. Cunha, Heckman, and Navarro (2005) and Cunha and Heckman (2016) use observations on individuals' decision of whether to attend college to help disentangle heterogeneity (known to individuals at the time of making the college attendance choice) from uncertainty (realized and observed later on in the life cycle). They estimate that 50 percent of the variation in lifetime earnings is due to heterogeneity. We consider lifetime earnings post schooling choice and estimate that between 65 percent and 84 percent is due to heterogeneity, depending on when in the life cycle we assume ε_0 to be observed by agents. Indeed, even though so far we have implicitly interpreted ε_0 as heterogeneity realized upon or before entry into the labor market, since ε_0 affects wages but does not affect any labor market decisions in our model, the exact timing of the realization of ε_0 is largely a matter of assumption. We would need to model additional choices, such as consumption/savings decisions, in order to parse out how much is heterogeneity known ex ante versus how much is uncertainty realized later in the life cycle.

³⁶ See online Appendix Section B.B3 for a formal presentation of this statement.

³⁷ Note that, unlike in the three-dimensional case, where we further transformed the first three principal components based on exclusion restrictions for interruptibility, here we keep the first principal component as it is produced

TABLE 6—FURTHER DECOMPOSITION OF $\text{var} \ln Q_{it}$

	Percent share of $\text{var} \ln Q_{it}$ due to...			
	x_0	x_{0C}	x_{0M}	x_{0I}
Whole sample	65.0	58.9	11.9	19.3
College+	17.2	9.4	2.8	3.6
Some college	27.5	22.2	10.6	8.0
Non-college	37.9	33.9	27.7	10.5

Notes: Level of experience: $t = 10$ years. The share of the variance explained by the individual skills x_{0k} do not sum to the share explained by x_0 due to the fact that (x_{0C}, x_{0M}, x_{0I}) are not mutually independent, and to nonlinearities.

TABLE 7—DECOMPOSITION OF $\text{var} \ln Q_{it}$: ONE-DIMENSIONAL MODEL

	Percent share of $\text{var} \ln Q_{it}$ due to...			
	Initial skills x_0 (term 1)	Shocks (term 2)	Heterogeneity ε_0 (term 3)	Education $ x_0$ (term 4)
Whole sample	32.5	3.94	60.4	3.16
College+	10.6	6.01	81.3	2.08
Some college	28.0	3.94	67.6	0.43
Non-college	24.6	3.73	71.6	0.13

Note: Level of experience: $t = 10$ years.

specific worker skills x instead of the three-dimensional $\mathbf{x} = (x_C, x_M, x_I)$ considered so far. We estimate the resulting “one-dimensional” version of our model and produce the decomposition of $\text{var} \ln Q_{it}$ in exactly the same way as we did for our full, three-dimensional model.³⁸

A comparison of Tables 7 and 5 immediately shows that, as might have been expected, the single-dimensional skill index model is capable of explaining substantially less of the variance of career output based on observables than our original model, which allows for three-dimensional skill bundles. This reduced explanatory power of the single-dimensional skill index manifests itself in two different ways. First, and most directly, the share of variance explained by observed initial worker skills (first column in Tables 7 and 5) tends to be smaller in the scalar case than in the three-dimensional case. Second, the share explained by labor market shocks (second column in Tables 7 and 5) is markedly smaller in the scalar case than in the three-dimensional case. Intuitively, two occupations that look similar in terms of the single-dimensional attribute y (which is constructed as the first principal component of our set of O*NET descriptors) may look rather different in terms of the three-dimensional bundle \mathbf{y} (which is based on the first three principal components), along dimensions of \mathbf{y} that are not captured by the single-dimensional index y .

by the PCA, and remain agnostic as to the label of the “skill” that this captured by this scalar measure. Details of this construction, and detailed estimation results for this one-dimensional version of our model are available on request.

³⁸In the notation of Section II, we redo the entire estimation exercise replacing $L = 3$ by $L = 1$. This results in a reduction in the number of data moments to be calculated (for example, the covariance terms involving y_C and y_M disappear), as well as a reduction in the number of parameters to be estimated (for example, there are fewer parameters describing the production function and the distribution of y).

What labor market shocks do is move workers between occupations with different productive attributes. If two occupations have a similar single-dimensional attribute y , but different three-dimensional bundles \mathbf{y} , then moving a worker between those two occupations will not cause a lot of variation in that worker's career output in the single-dimensional model, which fails to pick up the differences between the two occupations, but will do so in the three-dimensional world where the two occupations do look different. Similarly, two workers may look similar in terms on the single-dimensional attribute x , but very different along dimensions of \mathbf{x} that are not captured by x . When these workers are confronted with the opportunity of a job with skill requirements summarized by y , the scalar model will not generate any wage variation. The multidimensional model will generate substantial wage variation between these workers if one is much better suited than the other in terms of the vector of requirements \mathbf{y} .

While the model with unidimensional skill underestimates the share of the variance explained by both initial skills and shocks during working life, it underestimates the effect due to shocks more. Comparing the first rows of Tables 5 and 7, we see that the model with multidimensional skills implies initial skills explain 3.96 times as much of the variance as career shocks when looking at the entire sample. The corresponding ratio for the unidimensional model is 8.25, reducing the relative role of shocks by more than half. This same relative reduction of the role for shocks occurs within each education group.

Next we compute the cohort-wide average elasticities of expected career output (at ten years of experience) with respect to the subset of model parameters listed in the first column of Table 8. The way in which we construct some of those elasticities requires some explanation. The first three rows of Table 8 aim to measure the impact on average career output of a marginal proportional change in initial workforce skills. What we mean by that is the following. For each separate skill dimension k , we can multiply the initial endowment x_{0k} of every worker in the sample by a common scale $1 + \Delta$ and compute their career output $Q_{it}(\Delta)$ under this modified distribution of initial skills. The numbers reported in the first three rows of Table 8 are $E[\partial \ln Q_{it}(\Delta) / \partial \Delta]$, evaluated at $\Delta = 0$.

The following three rows turn to the labor demand side and assess the impact on average career output of a marginal change in the skill requirements demanded by employers. We measure that by the average elasticity of Q_{it} with respect to the mean skill requirement in each separate dimension.³⁹

The next three rows are about the matching technology, starting with the (self-explanatory) elasticities of Q_{it} with respect to the frequencies of job contacts, λ_0 and λ_1 , followed by a row that we label *Mismatch*. Numbers in this last row are based on counterfactual simulations where we reassign employed workers as follows. If a worker with current skills \mathbf{x} is employed in a job with attributes \mathbf{y} by the time they reach ten years of experience, we reassign them to a job of type $\mathbf{y}' = (1 - \Delta)\mathbf{y} + \Delta\mathbf{y}^*(\mathbf{x})$, where Δ is a small, positive number and where $\mathbf{y}^*(\mathbf{x})$

³⁹The marginal sampling distribution of skill requirement k is parameterized as a Beta distribution with parameters (η_k^1, η_k^2) . Those parameters obviously map 1:1 into the mean and variance of that marginal distribution. To compute the elasticity of Q_{it} with respect to the mean sampled skill requirement in dimension k , $E_T y_k$, we vary (η_k^1, η_k^2) to induce a marginal change in the mean, keeping the variance constant.

TABLE 8—ELASTICITIES OF Q_{it}

Elasticity of Q_{it} with respect to:	Whole sample	College+	Some college	Non-college
x_{0C}	0.31	0.33	0.32	0.28
x_{0M}	−0.11	−0.10	−0.13	−0.11
x_{0I}	0.07	0.07	0.06	0.06
$E_{\Upsilon y_C}$	0.08	0.13	0.09	0.04
$E_{\Upsilon y_M}$	0.07	0.06	0.07	0.07
$E_{\Upsilon y_I}$	0.06	0.05	0.06	0.06
λ_0	0.02	0.03	0.02	0.02
λ_1	0.01	0.01	0.04	0.02
Mismatch	0.26	0.31	0.26	0.22

Note: Level of experience: $t = 10$ years.

is the output-maximizing job type for a worker with skills \mathbf{x} . We can compute that worker's career output $Q_{it}(\Delta)$ following this reassignment for any value of Δ . The numbers reported in the *Mismatch* row of Table 8 $E[\partial \ln Q_{it}(\Delta)/\partial \Delta]$, evaluated at $\Delta = 0$.

Comparing the elasticities for the whole sample, we see that the largest effect on output is associated with a marginal improvement in the cognitive skills of the entering workforce. Interestingly, a marginal increase in the interpersonal skill endowment of entrants only has a comparatively small effect and an increase in their manual skills would have a negative effect, the latter being a result of the small effect of manual skills on output and the large effect on disutility of work.

The (quantitatively close) second largest effect on output is associated with a marginal reduction in the mismatch between worker skill and job requirements. This one-time reduction in mismatch has a direct effect on increasing current output, but also an effect on future output through improved skill accumulation of workers. The output loss due to mismatch dwarfs the losses associated with the sheer frequency of worker-firm contacts, λ_0 and λ_1 , which our model estimates to be an order of magnitude smaller in comparison.

Finally, increasing the mean skill requirement in the sampling distribution has moderately positive effects on output, which are roughly equal across skill types.

Looking across the education groups the same broad pattern holds as in the full sample. Reductions in mismatch and increases in the cognitive ability of the entering workforce have the largest effects on output. The magnitudes of both effects increase somewhat in education.

VII. Conclusion

We extend a standard and well-tested search-theoretic model of individual careers to allow for multidimensional skills and on-the-job learning. We estimate the model using occupation-level measures of skill requirements based on O*NET data, combined with a worker-level panel (NLSY79). We use the estimated model to shed light on the origins and costs of mismatch along three dimensions of skills: cognitive, manual, and interpersonal.

Our main findings are the following. The model sees cognitive, manual, and interpersonal skills as very different productive attributes. Manual skills have moderate

returns and adjust quickly (i.e., they are easily accumulated on the job, and relatively easily lost when left unused). Cognitive skills have much higher returns, but are much slower to adjust. The returns on interpersonal skills are similar to (slightly higher than) those on manual skills. Yet interpersonal skills are all but fixed over a worker's lifetime. Next, the cost of skill mismatch (modeled as the combination of an output loss and a loss of worker utility caused by skill mismatch) is very high for cognitive skills, an order of magnitude greater than for manual or interpersonal skills. Moreover, this cost is asymmetric: employing a worker who is under-qualified in cognitive skills (i.e., has a level of skills that falls short of the job's skill requirements) is several orders of magnitude more costly than employing an over-qualified worker. Those important differences between various skill dimensions are missed when subsuming worker productive heterogeneity into one single scalar index. This is highlighted when we decompose the variance of lifetime output between initial skills (observed and unobserved) and shocks over the career. The use of a unidimensional model of skill rather than a multidimensional model underestimates the explained share of the variance (of the present value of output) by half and, additionally, underestimates the contribution of shocks relative to initial skills by half.

APPENDIX A. MATHEMATICAL APPENDIX

A. Solving for the Value Functions

The value functions can be solved for in quasi-closed form. We first focus on the match value $P(\mathbf{x}, \mathbf{y})$, taking the value of unemployment $U(\mathbf{x})$ as given. To solve for $P(\mathbf{x}, \mathbf{y})$, it is convenient to parameterize P and \mathbf{x} as a function of the worker's tenure, say t , in the job under consideration. The solution to the first-order linear PDE (2) is then characterized by the following system of $K + 1$ ODEs:

$$(A1) \quad \frac{dx_k}{dt} = g_k(\mathbf{x}(t), \mathbf{y}), \quad k = 1, \dots, K,$$

$$(A2) \quad \frac{dz}{dt} = (r + \mu + \delta)z - [f(\mathbf{x}(t), \mathbf{y}) - c(\mathbf{x}(t), \mathbf{y})] - \delta U(\mathbf{x}(t)),$$

which are indeed the characteristic equations of (2). Match value is then the solution to $P(\mathbf{x}(t), \mathbf{y}) = z(t)$. Initial conditions for the first K equations (A1) are given by the worker's skill vector $\mathbf{x}(0)$ at the point of hire. The last initial condition, $z(0)$, is unknown, but we can impose the boundary condition $z(t)\exp[-(r + \mu + \delta)t] \rightarrow 0$ as $t \rightarrow +\infty$ to pin down a unique solution to (A2).

To be more explicit, let us denote by $\mathbf{X}(t; \mathbf{y}, \mathbf{x}_0)$ the solution to (A1) given initial condition \mathbf{x}_0 and job type \mathbf{y} (possibly equal to $\mathbf{0}_L$ if the worker is unemployed). The date- t value of a match between a job with attributes \mathbf{y} and a worker with current skill bundle $\mathbf{x}(t)$ is then given by the solution to (A2):

$$P(\mathbf{x}(t), \mathbf{y}) = \int_t^{+\infty} [f(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t)), \mathbf{y}) - c(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t)), \mathbf{y}) + \delta U(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t)))] e^{-(r+\mu+\delta)(s-t)} ds.$$

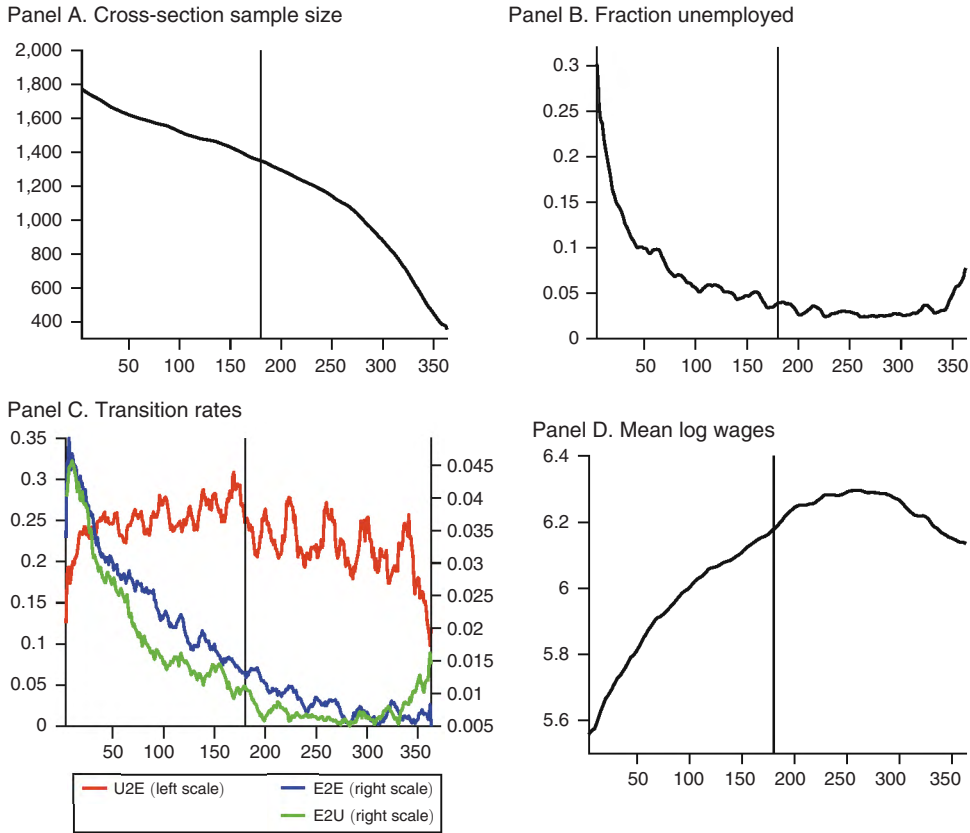


FIGURE A1. SAMPLE DESCRIPTION

The value of unemployment $U(\mathbf{x}(t)) = \int_t^{+\infty} b(\mathbf{X}(s; \mathbf{0}, \mathbf{x}(t))) e^{-(r+\mu)(s-t)} ds$ is solved for in a similar fashion, and the surplus associated with a typical match is obtained by subtraction:

$$\begin{aligned}
 (A3) \quad & P(\mathbf{x}(t), \mathbf{y}) - U(\mathbf{x}(t)) \\
 &= \int_t^{+\infty} \left[f(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t)), \mathbf{y}) - c(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t)), \mathbf{y}) - b(\mathbf{X}(s; \mathbf{y}, \mathbf{x}(t))) \right] \\
 &\quad \times e^{-(r+\mu+\delta)(s-t)} ds.
 \end{aligned}$$

Using the functional forms assumptions (8) and (9) in the formula above produces (10).

B. Data Details

Our final estimation sample consists of an initial cross-section of 1,840 males whom we follow over up to 30 years. There is, however, a substantial amount of attrition, which we comment on in the next paragraph.

TABLE A1—SUMMARY OF MODEL PARAMETERS

Structural object	Symbol	How it is estimated
Offer arrival rates	(λ_0, λ_1)	Estimated within the model
Job destruction rate	δ	Pre-estimated as sample mean E2U rate
Unemployment income	b	Estimated within the model
Production function f	$(\alpha_T, \alpha_C, \alpha_M, \alpha_I, \alpha_{CC}, \alpha_{MM}, \alpha_{II}, \kappa_C^U, \kappa_M^U, \kappa_I^U)$	Estimated within the model
Utility cost of over-qualification c	$(\kappa_C^O, \kappa_M^O, \kappa_I^O)$	Estimated within the model
Skill accumulation function g	$(\gamma_C^U, \gamma_M^U, \gamma_I^U, \gamma_C^O, \gamma_M^O, \gamma_I^O, g)$	Estimated within the model
Joint distribution of initial worker skills N		Directly observed from initial sample cross section
General worker efficiency x_T	$(\zeta_S, \zeta_C, \zeta_M, \zeta_I, \varepsilon_0)$	Distribution of ε_0 estimated by deconvolution in final step
Sampling distribution of job attributes Υ	$(\xi_C, \xi_M, \xi_I, \rho_{CM}, \rho_{CI}, \rho_{MI}, \eta_C^1, \eta_C^2, \eta_M^1, \eta_M^2, \eta_I^1, \eta_I^2)$	Estimated within the model, specified as Gaussian Copula with Beta marginals
Attrition and discount rates	(r, μ)	Calibrated outside model

Flows, Stocks, and Wages over Time.—Figure A1 describes our sample in terms of a set of times series about worker stocks, labor market transition rates, and average wages over the full 30-year sample window. The horizontal-axis variable is time, measured in months since labor market entry.

Panel A of Figure A1 shows the pattern of attrition from our sample. Attrition is initially very gradual, with the sample cross-section size declining by about 30 percent over the initial 20 years. Past that point, attrition accelerates considerably. This is partly a consequence of the fact that we follow a cohort of individuals from the date they leave full-time education, resetting time to zero on the week they enter the labor market. Individuals having spent more time at school enter the labor market later, and are therefore observed for fewer years than less educated individuals. This causes the composition of the sample to shift toward less educated individuals as one approaches the end of the observation window. To circumvent this problem, we restrict our estimation sample to the first 15 years (180 months) of the initial sample. This 180-month cutoff is materialized by a thick vertical black line on all panels of Figure A1.

Panel B shows the non-employment rate among sample members. As one would expect, this rate declines monotonically over time, until it reaches a steady level slightly under 5 percent. It rises again slightly after about 20/25 years, likely as a result of the compositional shift discussed above. Perhaps slightly more surprising is the long time it takes for the non-employment rate to reach this steady state (roughly ten years). Panel C shows the rates of transition between labor market states. The non-employment exit rate is roughly stable at around 25 percent per month, while the transition rates from job to job and into non-employment decline smoothly over the sample window. Finally, panel D plots average log wages among employed sample members which, again as one would expect, increase monotonically over time until they reach a point where, mirroring the non-employment rate, they start declining, again a likely consequence of nonrandom attrition from the sample. Wages are converted to real 1983 dollars using the regional all items CPI (Bureau of Labor Statistics 1978-2012).

TABLE A2—EXAMPLES OF SKILL REQUIREMENT SCORES

Occupation title	Skill requirements		
	Cognitive	Manual	Interpersonal
Physicists	1	0.755	0.692
Graders and sorters, agricultural products	0	0.138	0.058
Aircraft mechanics and service technicians	0.613	1	0.318
Telemarketers	0.147	0	0.330
Preventive medicine physicians	0.658	0.410	1
Molding, coremaking, and casting machine setters, operators, and tenders, metal and plastic	0.302	0.641	0

Sources: O*NET and authors' calculations

Worker Skills and Job Skill Requirements.—Table A2 lists some examples of the cognitive, manual, and interpersonal skill requirement scores we constructed for a few occupations. We denote those scores by $\tilde{\mathbf{y}} = (\tilde{y}_C, \tilde{y}_M, \tilde{y}_I)$ and will use them as empirical measures of the model's job attributes \mathbf{y} . Examples in Table A2 include the occupations with the highest cognitive (Physicist), manual (Aircraft Mechanics and Service Technicians), and interpersonal (Preventive Medicine Physicians) skill requirements in the sample, and the occupations with the lowest cognitive (Graders and Sorters, Agricultural Products), manual (Telemarketers), and interpersonal (Molding, Coremaking, and Casting Machine Setters, Operators, and Tenders, Metal and Plastic) skill requirements.

The correlation pattern of workers' *initial* skills and the skill requirements of the *first jobs they are observed in* is described in Table A3, where workers' initial cognitive, manual, and interpersonal skill indices are denoted by (x_{C0}, x_{M0}, x_{I0}) , while $(\tilde{y}_C, \tilde{y}_M, \tilde{y}_I)$ refer to the empirical measures of job skill requirements in a worker's first job. This correlation pattern reveals several features of the data. First, (x_{C0}, x_{M0}, x_{I0}) are positively correlated in our cross section of workers. Even though those correlation coefficients are far below 1, they suggest that workers with high skills in one dimension tend to have high skills in the other two. Cognitive and manual skills appear slightly more strongly associated with each other than either is with interpersonal skills. Second, \tilde{y}_C is positively correlated with both \tilde{y}_M and \tilde{y}_I in the cross section of workers' first jobs. Even though there is obviously some selection here (as the set of jobs a worker will take up depends on their own skill bundle \mathbf{x}), this suggests that jobs requiring high levels of cognitive skills also tend to require high skill levels in one of the manual or interpersonal dimensions. While manual and interpersonal skill requirements are both positively correlated with cognitive skill requirements, they are negatively correlated with each other. Third, (x_{C0}, \tilde{y}_C) , (x_{M0}, \tilde{y}_M) , and (x_{I0}, \tilde{y}_I) are positively correlated (as expected), and so are (x_{C0}, \tilde{y}_I) , (x_{M0}, \tilde{y}_C) , (x_{I0}, \tilde{y}_M) , (x_{I0}, \tilde{y}_C) . By contrast, (x_{C0}, \tilde{y}_M) and (x_{M0}, \tilde{y}_I) are negatively correlated, suggesting that workers select themselves into either manual or non-manual jobs, as fits their skill bundles.

Occupation-Specific Wage Deciles.—In columns 4–6 of Table 3 we control for a worker's position within their occupation-specific wage distribution. Specifically, we include an indicator variable for which decile of the occupation-specific distribution each worker belongs to. We construct the cutoffs for the deciles using the 1980

TABLE A3—CORRELATION PATTERN OF INITIAL SKILLS
AND SKILL REQUIREMENTS IN FIRST JOB

	x_{C0}	x_{M0}	x_{I0}	\tilde{y}_C	\tilde{y}_M	\tilde{y}_I
x_{C0}	1					
x_{M0}	0.46	1				
x_{I0}	0.39	0.31	1			
\tilde{y}_C	0.48	0.15	0.23	1		
\tilde{y}_M	-0.18	0.18	0.01	0.32	1	
\tilde{y}_I	0.55	-0.01	0.27	0.64	-0.34	1

Census (Ruggles et al. 2019). From the Census, we restrict the sample to include men, aged 16 to 65 who worked at least 26 weeks in the year. We convert weekly earnings to real 1982 dollars using the CPI to be comparable with the NLSY data. We map the 1980 occupation codes to the 1970 occupation using Table 1 from Census Bureau (1989).

C. Construction of Skill Measures

The datasets from which we construct worker skill and job skill requirement scores both consist of a set of P different measures observed for N individuals (workers in the case of the NLSY, and occupations in the case of O*NET). We denote the $N \times P$ matrix of all observations by \mathbf{M} . PCA decomposes the matrix \mathbf{M} as $\mathbf{M} = \mathbf{F}\mathbf{L}$, where \mathbf{F} is the orthonormal $N \times P$ matrix of principal eigenvectors of $\mathbf{M}^\top \mathbf{M}$ and \mathbf{L} is a $P \times P$ matrix of factor loadings. We consider the first 3 principal components only, i.e., we consider the decomposition $\mathbf{M} = \mathbf{F}_3 \mathbf{L}_3 + \mathbf{U}$, where \mathbf{F}_3 is the $N \times 3$ matrix formed by taking the first 3 columns of \mathbf{F} and \mathbf{L}_3 is the $3 \times P$ matrix formed by taking the first 3 rows of \mathbf{L} .

For any invertible 3×3 matrix \mathbf{T} , the above decomposition of \mathbf{M} can be rewritten as $\mathbf{M} = (\mathbf{F}_3 \mathbf{T})(\mathbf{T}^{-1} \mathbf{L}_3) + \mathbf{U}$, which is an alternative decomposition of \mathbf{M} into new (linearly recombined) factors $\mathbf{F}_3 \mathbf{T}$ with loadings $\mathbf{T}^{-1} \mathbf{L}_3$. We choose \mathbf{T} such that our decomposition of \mathbf{M} satisfies our chosen exclusion restrictions. Taking the case of O*NET as an example, we order the measures such that measure 1 (the first column of \mathbf{M}) is the score on *mathematics knowledge*, measure 2 is the score on *mechanical knowledge*, and measure 3 is the score on *social perceptiveness*, then define $\mathbf{T} = \mathbf{L}_{3,3}$ where $\mathbf{L}_{3,3}$ is the 3×3 matrix made up of the first three columns of \mathbf{L}_3 . It should be emphasized that our method of constructing worker skill and job skill requirement scores differs slightly from the approach usually taken in the related literature. The conventional approach consists of assigning each of the P data measures (of skills or skill requirements, as the case may be) to one of K different bins, where K is the number of skill dimensions relevant to the model (three, in our case), and set the score in skill dimension k as the average of all measures in bin k .

The conventional method therefore assumes that any given measure is only relevant to one single skill dimension. Which skill dimension a measure is relevant to must be decided a priori. In our case, this would mean deciding for every NLSY or O*NET descriptor whether it relates to cognitive, manual, or interpersonal skills. While this decision may seem relatively straightforward, at least on an intuitive

TABLE A4—ALTERNATIVE EXCLUSION RESTRICTIONS IN O*NET

Cognitive skill requirements	Mathematics knowledge
Mathematical reasoning	0.983 (0.009)
Fluency of ideas	0.930 (0.017)
Written comprehension	0.918 (0.019)
Oral comprehension	0.902 (0.020)
Manual skill requirements	Mechanics knowledge
Finger dexterity	0.973 (0.011)
Repairing and maintaining mechanical equipment	0.940 (0.016)
Arm-hand steadiness	0.717 (0.033)
Manual dexterity	0.657 (0.036)
Interpersonal skill requirements	Social perceptiveness
Selling or influencing others	0.997 (0.004)
Negotiation	0.991 (0.006)
Persuasion	0.979 (0.010)
Speaking	0.965 (0.012)

Notes: Correlation with alternative exclusion restrictions. Standard errors in parentheses.

level, for some measures (for instance the six measures on which we impose exclusion restrictions), it is far from clear-cut for most measures, which can easily be argued to be relevant for two or more skill dimensions. We therefore choose to minimize the number of exclusion restrictions we impose on the data. We believe that our approach offers a good compromise between interpretability, parsimony, and ability to capture the covariance patterns between skills and skill requirement dimensions.

*D. Robustness to Alternative Exclusion Restrictions in O*NET*

In Table A4 we present the correlation between our preferred skill requirement measures and skill requirement measures constructed using four alternative exclusion restrictions for each of the cognitive, manual, and interpersonal dimensions. Specifically, we consider Mathematical Reasoning, Fluency of Ideas, Written Comprehension, and Oral Comprehension as alternative exclusion restrictions to Mathematics Knowledge to anchor the cognitive skill requirements. We consider Finger Dexterity, Repairing and Maintaining Mechanical Equipment, Arm-Hand Steadiness, and Manual Dexterity as alternatives to Mechanics Knowledge to

anchor the manual skill requirements. We consider Selling or Influencing Others, Negotiation, Persuasion, and Speaking as alternative exclusion restrictions to Social Perceptiveness to anchor interpersonal skill requirements. The correlations are all very high, especially for the alternatives we consider for cognitive and interpersonal anchors. For the manual anchor, the correlation is very high for the Finger Dexterity and Repairing and Maintaining Mechanical Equipment alternatives, but somewhat weaker for both the Arm-Hand Steadiness and Manual Dexterity alternatives. As a result it is important to keep in mind that our notion of manual skill requirements has more to do with the ability to fix machines than the ability to do very physical labor. This aligns well with our exclusion for the worker skill side in the NLSY79 which measures auto mechanics and machine shop knowledge.

E. Identification

This Appendix contains a formal discussion of identification. Identification is, in large part, parametric, in that many of the arguments below make use of the specific functional forms assumed in the main text.

The job loss rate δ is directly observed in the data. We assume that so is the population distribution of initial skill bundles. Moreover, we discuss identification conditional on knowledge of the discount rate r and the sample attrition rate μ .

The wage equation (5) can be written as

$$\begin{aligned}
 (A4) \quad w(\mathbf{x}, \mathbf{y}, \sigma) = & \sigma f(\mathbf{x}, \mathbf{y}) + (1 - \sigma)b(\mathbf{x}) + (1 - \sigma)c(\mathbf{x}, \mathbf{y}) \\
 & - \lambda_1(1 - \sigma)[P(\mathbf{x}, \mathbf{y}) - U] \int_{\mathbf{y}'} \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq P(\mathbf{x}, \mathbf{y})\} d\Upsilon(\mathbf{y}') \\
 & - \lambda_1 \int_{\mathbf{y}'} [\mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq \sigma P(\mathbf{x}, \mathbf{y}) + (1 - \sigma)U\} \\
 & \quad - \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq P(\mathbf{x}, \mathbf{y})\}] \\
 & \times [P(\mathbf{x}, \mathbf{y}') - \sigma P(\mathbf{x}, \mathbf{y}) - (1 - \sigma)U] d\Upsilon(\mathbf{y}').
 \end{aligned}$$

A first important implication of (A4) is that the maximum wage given (\mathbf{x}, \mathbf{y}) is $f(\mathbf{x}, \mathbf{y})$, implying in turn that the maximum wage given \mathbf{y} is $f(\mathbf{y}, \mathbf{y}) = x_T \cdot (\alpha_T + \sum_{k=C,M,I} (\alpha_k y_k + \alpha_{kk} x_k \cdot y_k))$. Because \mathbf{y} is observed for all employed workers,⁴⁰ and because x_T is also a function of observables (up to the uncorrelated heterogeneity term ε_0), namely the worker's education, initial skill bundle and experience, this proves identification of the parameters $\alpha_T, \alpha_C, \alpha_M, \alpha_I, \alpha_{CC}, \alpha_{MM}, \alpha_{II}, g, \zeta_S, \zeta_C, \zeta_M, \zeta_I$.

Next, consider the set of workers with initial skill bundle \mathbf{x} exiting non-employment at any experience level. The (observed) set of job types \mathbf{y} that those workers accept is the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) \geq U\}$, and its boundary is the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) = U\}$. This latter set is therefore identified, conditional on knowledge of \mathbf{x} . We now show that

⁴⁰What is, in fact, observed, is not directly \mathbf{y} but rather its empirical counterpart $\tilde{\mathbf{y}}$. With our functional form assumptions on $f(\mathbf{y}, \mathbf{y})$ and $y_k = \tilde{y}_k^{\zeta_k}$, $k = C, M, I$, the maximum wage given \mathbf{y} jointly identifies the α s and the ζ s.

this latter fact allows identification of the parameters of the match value function $P(\mathbf{x}, \mathbf{y}) = U$.

First, from the expression of the match surplus (10), one can show that joint observation of \mathbf{x} and the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) = U\}$ allows separate identification of the parameters of $P(\mathbf{x}, \mathbf{y})$, i.e., the composite parameters $\kappa_k^{u/o} / (r + \delta + \mu - g + 2\gamma_k^{u/o})$, $k = C, M, I$.⁴¹ Now, the issue is that we do not directly observe worker skills at all levels of experience: rather, we only observe workers' *initial* skill bundles. However, consider a worker with (observed) initial skill bundle $\mathbf{x}(0)$ starting his working life in unemployment, and who finds a job after an initial unemployment spell of duration $d^{(1)}$. From the human capital accumulation function (6), we know that this worker's skill bundle by the time s/he finds a job is $\mathbf{x}(d^{(1)}) = (x_C(0)e^{-\gamma_C^o d^{(1)}}, x_M(0)e^{-\gamma_M^o d^{(1)}}, x_I(0)e^{-\gamma_I^o d^{(1)}})$. Identification of the parameters of $P(\mathbf{x}, \mathbf{y})$, $\kappa_k^{u/o} / (r + \delta + \mu - g + 2\gamma_k^{u/o})$, is thus obtained from the set of initially unemployed workers whose initial unemployment spell duration $d^{(1)} \rightarrow 0$. Furthermore, once the parameters of $P(\mathbf{x}, \mathbf{y})$ are known, observation of the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) = U\}$ given $\mathbf{x} = (x_C(0)e^{-\gamma_C^o d^{(1)}}, x_M(0)e^{-\gamma_M^o d^{(1)}}, x_I(0)e^{-\gamma_I^o d^{(1)}})$ with $\mathbf{x}(0)$ observed identifies γ_k^o for $k = C, M, I$. Combining those results, we now have separate identification of γ_k^o , κ_k^o , and $\kappa_k^u / (r + \delta + \mu - g + 2\gamma_k^u)$, and still need to separate κ_k^u from γ_k^u in the latter composite parameter. This can be done by repeating the latter argument for workers who are initially employed in matches with skill requirements $\mathbf{y}^{(1)}$ for which they are *under-qualified*, i.e., such that $x_k(0) < y_k^{(1)}$ for $k = C, M, I$, become unemployed after an initial spell duration of $d^{(1)}$, then find a job again after an unemployment spell of duration $d^{(2)}$. From the human capital accumulation function (6), those workers' skill bundles when they find their second job (at experience $d^{(1)} + d^{(2)}$) is given by $x_k(d^{(1)} + d^{(2)}) = e^{-\gamma_k^u d^{(2)}} [y_k^{(1)} - e^{-\gamma_k^u d^{(1)}} (y_k^{(1)} - x_k(0))]$. The only unknown parameter in this expression is γ_k^u , which is then again identified from the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) = U\}$.

The full set of production, utility, and human capital accumulation parameters is thus identified. Note that, while the arguments laid out above rely on the specific functional forms assumed in the main text, the background source of identification for the cost of mismatch and the speed of human capital accumulation (or decay) is a comparison of the set of job types \mathbf{y} that are acceptable to workers with equal initial skills $\mathbf{x}(0)$, but have experienced different employment histories.

Once the parameters of the human capital accumulation function $\mathbf{g}(\mathbf{x}, \mathbf{y})$ are known, we can construct any worker's full path of skill bundles \mathbf{x} : consider a worker in his n th spell (which could be a spell of unemployment). Denote the skill requirements in that spell by $\mathbf{y}^{(n)} = (y_C^{(n)}, y_M^{(n)}, y_I^{(n)})$ (both equal to 0 if the spell is one of unemployment), the worker's skill bundle at the beginning of that spell by $\mathbf{x}^{(n)} = (x_C^{(n)}, x_M^{(n)}, x_I^{(n)})$, and the duration of that spell by $d^{(n)}$. Spell duration $d^{(n)}$

⁴¹ One way to see this is to realize from (10) that the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) = U\}$ is the union of four quarter-ellipses, the centers and axes of which can be expressed as simple functions of \mathbf{x} and the parameter combinations $\kappa_k^{u/o} / (r + \delta + \mu - g + 2\gamma_k^{u/o})$. Observation of \mathbf{x} and \mathbf{y} for this set identifies these centers and axes.

and the vector $\mathbf{y}^{(n)}$ are observed in all spells, while $\mathbf{x}^{(n)}$ is only observed in the initial spell, $n = 1$, where it equals $\mathbf{x}(0)$. Then, using the skill accumulation equation $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, \mathbf{y})$, we have that $\mathbf{x}^{(n+1)} = \mathbf{X}(d^{(n)}; \mathbf{y}^{(n+1)}, \mathbf{x}^{(n)})$, where $\mathbf{X}(\cdot)$ denotes the solution to (A1) as explained in the main text. Using backward substitution, we can then construct $\mathbf{x}^{(n)}$ for any spell as a function of the history of durations and skill requirements of past spells and the worker's initial skill level $\mathbf{x}^{(1)} = \mathbf{x}(0)$.

Next, the set of job offers accepted by unemployed workers with skills \mathbf{x} identifies the sampling distribution $\Upsilon(\mathbf{y})$ over the set $\{\mathbf{y} : P(\mathbf{x}, \mathbf{y}) \geq U\}$. Thus, $\Upsilon(\mathbf{y})$ is (nonparametrically, conditionally on the rest of the model) identified over the union of all such sets for all skill bundles \mathbf{x} observed in the sample. That is, $\Upsilon(\mathbf{y})$ is identified at all skill requirement levels \mathbf{y} that are acceptable by at least some worker types.

Finally, the offer arrival rates λ_0 and λ_1 are identified, conditionally on the rest of the model, from sample U2E and E2E transition probabilities, and the flow value of non-employment, $b(\mathbf{x})$, is identified from the wage of workers exiting non-employment: applying (A4) to workers just exiting non-employment ($\sigma = 0$) yields

$$\begin{aligned} w(\mathbf{x}, \mathbf{y}, 0) &= b(\mathbf{x}) + c(\mathbf{x}, \mathbf{y}) \\ &+ \lambda_1 \int_{\mathbf{y}'} \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq P(\mathbf{x}, \mathbf{y})\} [P(\mathbf{x}, \mathbf{y}') - P(\mathbf{x}, \mathbf{y})] d\Upsilon(\mathbf{y}') \\ &- \lambda_1 \int_{\mathbf{y}'} \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq U\} [P(\mathbf{x}, \mathbf{y}') - U] d\Upsilon(\mathbf{y}'), \end{aligned}$$

which equals $b(\mathbf{x}) + c(\mathbf{x}, \mathbf{y})$ on the (known) set of \mathbf{y} s such that $P(\mathbf{x}, \mathbf{y}) = U$.

F. Unobserved Heterogeneity ε_0

The distribution of unobserved permanent worker heterogeneity ε_0 can be identified from the comparison of the simulated cross-section distribution of log wages *assuming away unobserved worker heterogeneity* (i.e., setting $\varepsilon_0 = 0$ for all workers) with the empirical distribution of log wages. For example, the variance of ε_0 can be estimated as the difference between the empirical wage variance and the variance of the simulated log-wage distribution without unobserved worker heterogeneity. We apply this strategy and compare the variances of the cross-section distributions of individual-level mean wages. By taking individual-level averages over time, we hope to minimize pollution of our estimate of $\text{var } \varepsilon_0$ by potential measurement error in the empirical wage data. The variance of $\ln \varepsilon_0$ thus estimated is 0.104 and accounts for 4.5 percent of the overall variance of individual-level mean wages.

G. Computation of Expected Career Output

The expected career output of a worker with current skills \mathbf{x} , employed in a job with requirements \mathbf{y} defined in (12) can be equivalently defined in the following recursive way:

$$\begin{aligned} \text{(A5)} \quad (r + \delta + \mu)Q(\mathbf{x}, \mathbf{y}) &= f(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) + \delta V(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} Q(\mathbf{x}, \mathbf{y}) \\ &+ \lambda_1 \int_{\mathbf{y}'} \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq P(\mathbf{x}, \mathbf{y})\} [Q(\mathbf{x}, \mathbf{y}') - Q(\mathbf{x}, \mathbf{y})] d\Upsilon(\mathbf{y}'), \end{aligned}$$

where the dependence of Q on fixed attributes (\mathbf{x}_0 , education, ε_0) is omitted and where $V(\mathbf{x})$ is the expected career output of an unemployed worker with current skill bundle \mathbf{x} . The latter is in turn defined by

$$(A6) \quad (r + \mu)V(\mathbf{x}) = b(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \mathbf{0}) \cdot \nabla V(\mathbf{x}) \\ + \lambda_0 \int \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq U(\mathbf{x})\} [Q(\mathbf{x}, \mathbf{y}') - V(\mathbf{x})] d\Upsilon(\mathbf{y}').$$

In those two formulas, $P(\mathbf{x}, \mathbf{y})$ and $U(\mathbf{x})$ denote the equilibrium (private) values of, respectively, a match between a type- \mathbf{x} worker and a type- \mathbf{y} job, and an unemployed type- \mathbf{x} worker. In both (A5) and (A6), the last (expectation) term captures the expected continuation value of the worker's career in his future matches. The social values $Q(\mathbf{x}, \mathbf{y})$ and $V(\mathbf{x})$ differ from the private values $P(\mathbf{x}, \mathbf{y})$ and $U(\mathbf{x})$ precisely because of those expectation terms: under our rent-sharing protocol, existing firm-worker pairs fail to internalize the value generated by the worker in future matches because all of said value is captured by the worker's future employer when he switches jobs. While the economic interpretation of those expectation terms is clear enough, mathematically their impact is to add a nonlinear term to the PDEs defining $Q(\mathbf{x}, \mathbf{y})$ and $V(\mathbf{x})$, which rules out any closed-form solution. Those equations must therefore be solved numerically, using the following procedure.

We choose a grid $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \times \{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ of n worker skill and m job skill requirement vectors. For any point $(\mathbf{x}_i, \mathbf{y}_j)$ on that grid, we approximate $Q(\mathbf{x}_i, \mathbf{y}_j) = \Pi(\mathbf{x}_i) \cdot \phi_Q(\mathbf{y}_j)$ and $V(\mathbf{x}_i) = \Pi(\mathbf{x}_i) \cdot \phi_V$, where $\Pi(\cdot)$ is a basis of complete polynomials of some chosen order, and where ϕ_V and $\phi_Q(\cdot)$ are a set of $m + 1$ vectors of coefficients that are computed by minimizing the distance between the left- and right-hand sides of (A5) and (A6) over the grid. Then, for a generic pair (\mathbf{x}, \mathbf{y}) that is not on the grid, we use the approximations $V(\mathbf{x}) = \Pi(\mathbf{x}) \cdot \phi_V$ and $Q(\mathbf{x}, \mathbf{y}) = \Pi(\mathbf{x}) \cdot \tilde{\phi}_Q(\mathbf{y})$, where $\tilde{\phi}_Q(\mathbf{y})$ is a linear interpolation of $\phi_Q(\tilde{\mathbf{y}}_{j1}(\mathbf{y}))$ and $\phi_Q(\tilde{\mathbf{y}}_{j2}(\mathbf{y}))$, $\tilde{\mathbf{y}}_{j1}(\mathbf{y})$ and $\tilde{\mathbf{y}}_{j2}(\mathbf{y})$ being the nearest two neighbors of \mathbf{y} on the grid.

Solving for ϕ_V and $\phi_Q(\cdot)$ involves repeated calculations of three-dimensional integrals like $\int \mathbf{1}\{P(\mathbf{x}, \mathbf{y}') \geq P(\mathbf{x}, \mathbf{y})\} [Q(\mathbf{x}, \mathbf{y}') - Q(\mathbf{x}, \mathbf{y})] d\Upsilon(\mathbf{y}')$, the computational cost of which increases very quickly as one increases the order of the polynomial basis $\Pi(\cdot)$ or the size (n, m) of the grid of points upon which the approximation is based. Fortunately, we only need to perform those calculations a limited number of times. In practice, we choose $m = n = 15$ and an approximation order of 3. We find that increasing the fineness of the approximation beyond those values makes little difference to the results.

REFERENCES

- Acemoglu, Daron, and David Autor.** 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1043–1171. Amsterdam: Elsevier.
- Anderson, Axel, and Lones Smith.** 2010. "Dynamic Matching and Evolving Reputations." *Review of Economic Studies* 77 (1): 3–29.
- Antonovics, Kate, and Limor Golan.** 2012. "Experimentation and Job Choice." *Journal of Labor Economics* 30 (2): 333–66.

- Autor, David H., and David Dorn.** 2013a. "Data for: The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97. <https://www.ddorn.net/data/Autor-Dorn-LowSkillServices-FileArchive.zip> (accessed June 18, 2019).
- Autor, David H., and David Dorn.** 2013b. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97.
- Autor, David H., Frank Levy, and Richard J. Murnane.** 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–1333.
- Ben-Porath, Yoram.** 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–65.
- Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel.** 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43 (4): 972–1059.
- Bureau of Labor Statistics.** 1978–2012. "Consumer Price Index - All Urban Consumers, All Item: USA - CUUR0000SA0; Northeast - CUUR0100SA0; Midwest - CUUR0200SA0; South - CUUR0300SA0; West - CUUR0400SA0." United States Department of Labor. <https://data.bls.gov/cgi-bin/surveymost?cu> (accessed September 25, 2014).
- Bureau of Labor Statistics.** 1979–2011. "National Longitudinal Survey of Youth 1979 cohort." United States Department of Labor, produced and distributed by the Center for Human Resource Research (CHRR), The Ohio State University. Columbus, OH. <https://www.bls.gov/nls/nlsy79.htm> (accessed September 12, 2014).
- Bureau of the Census.** 1989. "The Relationship between the 1970 and the 1980 Industry and Occupation Classification Systems." Technical Paper 59, <https://www2.census.gov/library/publications/1989/demographics/tech-paper-59.pdf>.
- Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin.** 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2): 323–64.
- Chiappori, Pierre-André, and Bernard Salanié.** 2016. "The Econometrics of Matching Models." *Journal of Economic Literature* 54 (3): 832–61.
- Cunha, Flávio, and James Heckman.** 2016. "Decomposing Trends in Inequality in Earnings into Forecastable and Uncertain Components." *Journal of Labor Economics* 34 (2): S31–S65.
- Cunha, Flavio, James Heckman, and Salvador Navarro.** 2005. "Separating Uncertainty from Heterogeneity in Life Cycle Earnings." *Oxford Economic Papers* 57 (2): 191–261.
- Dey, Matthew S., and Christopher J. Flinn.** 2005. "An Equilibrium Model of Health Insurance Provision and Wage Determination." *Econometrica* 73 (2): 571–627.
- Eeckhout, Jan, and Philipp Kircher.** 2011. "Identifying Sorting: In Theory." *Review of Economic Studies* 78 (3): 872–906.
- Ekeland, Ivar, James J. Heckman, and Lars Nesheim.** 2004. "Identification and Estimation of Hedonic Models." *Journal of Political Economy* 112 (S1): S60–S109.
- Gathmann, Christina, and Uta Schönberg.** 2010. "How General Is Human Capital? A Task-Based Approach." *Journal of Labor Economics* 28 (1): 1–49.
- Groes, Fane, Philipp Kircher, and Iouri Manovskii.** 2015. "The U-Shapes of Occupational Mobility." *Review of Economic Studies* 82 (2): 659–92.
- Guvener, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer.** 2020. "Multidimensional Skill Mismatch." *American Economic Journal: Macroeconomics* 12 (1): 210–44.
- Hagedorn, Marcus, Tzuo Hann Law, and Iouri Manovskii.** 2017. "Identifying Equilibrium Models of Labor Market Sorting." *Econometrica* 85 (1): 29–65.
- Heckman, James J., Rosa L. Matzkin, and Lars Nesheim.** 2010. "Nonparametric Identification and Estimation of Nonadditive Hedonic Models." *Econometrica* 78(5): 1569–91.
- Heckman, James J., and Guilherme Sedlacek.** 1985. "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy* 93 (6): 1077–1125.
- Heckman, James J., Jora Stixrud, and Sergio Urzua.** 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–82.
- Jovanovic, Boyan.** 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5): 972–90.
- Kambourov, Gueorgui, and Iouri Manovskii.** 2009a. "Occupational Mobility and Wage Inequality." *Review of Economic Studies* 76 (2): 731–59.
- Kambourov, Gueorgui, and Iouri Manovskii.** 2009b. "Occupational Specificity of Human Capital." *International Economic Review* 50 (1): 63–115.
- Keane, Michael P., and Kenneth I. Wolpin.** 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3): 473–522.

- Lazear, Edward P.** 2009. "Firm-Specific Human Capital: A Skill-Weights Approach." *Journal of Political Economy* 117 (5): 914–40.
- Lee, Donghoon, and Kenneth I. Wolpin.** 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica* 74 (1): 1–46.
- Lindenlaub, Ilse.** 2017. "Sorting Multidimensional Types: Theory and Application." *Review of Economic Studies* 84 (2): 718–89.
- Lindenlaub, Ilse, and Fabien Postel-Vinay.** 2020. "Multidimensional Sorting under Random Search." Unpublished.
- Lise, Jeremy, Costas Meghir, and Jean-Marc Robin.** 2016. "Matching, Sorting, and Wages." *Review of Economic Dynamics* 19: 63–87.
- Lise, Jeremy, and Fabien Postel-Vinay.** 2020. "Replication Data for: Multidimensional Skills, Sorting, and Human Capital Accumulation." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E117662v1>.
- Lise, Jeremy, and Jean-Marc Robin.** 2017. "The Macrodynamics of Sorting between Workers and Firms." *American Economic Review* 107 (4): 1104–35.
- Lucas, Robert E., and Edward C. Prescott.** 1974. "Equilibrium Search and Unemployment." *Journal of Economic Theory* 7 (2): 188–209.
- Moscarini, Giuseppe.** 2001. "Excess Worker Reallocation." *Review of Economic Studies* 68 (3): 593–612.
- National Center for O*NET Development.** 2014. "O*NET 18.1 Database, March 2014." Raleigh, NC. https://www.onetcenter.org/db_releases.html (accessed September 20, 2014).
- Neal, Derek.** 1999. "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics* 17 (2): 237–61.
- Pavan, Ronni.** 2011. "Career Choice and Wage Growth." *Journal of Labor Economics* 29 (3): 549–87.
- Poletaev, Maxim, and Chris Robinson.** 2008. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Survey, 1984–2000." *Journal of Labor Economics* 26 (3): 387–420.
- Postel-Vinay, Fabien, and Jean-Marc Robin.** 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6): 2295–2350.
- Rosen, Sherwin.** 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (1): 34–55.
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek.** 2019. "IPUMS USA: Version 9.0 1980 5% State Sample." Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D010.V9.0>. (accessed June 14, 2019).
- Sanders, Carl.** 2012. "Skill Uncertainty, Skill Accumulation, and Occupational Choice." https://economicdynamics.org/meetpapers/2012/paper_633.pdf.
- Sanders, Carl, and Christopher Taber.** 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annual Review of Economics* 4: 399–425.
- Shimer, Robert, and Lones Smith.** 2000. "Assortative Matching and Search." *Econometrica* 68 (2): 343–69.
- Sullivan, Paul.** 2010. "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search." *International Economic Review* 51 (1): 289–317.
- Taber, Christopher, and Rune Vejlin.** 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica* 88 (3): 1031–1069.
- Tinbergen, Jan.** 1956. "On the Theory of Income Distribution." *Weltwirtschaftliches Archiv* 77: 155–75.
- Yamaguchi, Shintaro.** 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30 (1): 1–53.