



Management Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Demand-Aware Career Path Recommendations: A Reinforcement Learning Approach

Marios Kokkodis, Panagiotis G. Ipeirotis

To cite this article:

Marios Kokkodis, Panagiotis G. Ipeirotis (2021) Demand-Aware Career Path Recommendations: A Reinforcement Learning Approach. Management Science 67(7):4362-4383. <https://doi.org/10.1287/mnsc.2020.3727>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2020, INFORMS

Please scroll down for article—it is on subsequent pages





With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Demand-Aware Career Path Recommendations: A Reinforcement Learning Approach

Marios Kokkodis,^a Panagiotis G. Ipeirotis^b

^a Carroll School of Management, Boston College, Chestnut Hill, Massachusetts 02467; ^b Leonard N. Stern School of Business, New York University, New York, New York 10012

Contact: kokkodis@bc.edu,  <https://orcid.org/0000-0002-5037-6060> (MK); panos@stern.nyu.edu,  <https://orcid.org/0000-0002-2966-7402> (PGI)

Received: October 8, 2017

Revised: February 6, 2019; January 3, 2020; May 6, 2020

Accepted: June 3, 2020

Published Online in Articles in Advance: October 8, 2020

<https://doi.org/10.1287/mnsc.2020.3727>

Copyright: © 2020 INFORMS

Abstract. A skill's value depends on dynamic market conditions. To remain marketable, contractors need to keep reskilling themselves continuously. But choosing new skills to learn is an inherently hard task: Contractors have very little information about current and future market conditions, which often results in poor learning choices. Recommendation frameworks could reduce uncertainty in learning choices. However, conventional approaches would likely be inefficient; they would model previous (often poor) observed contractor learning behaviors to provide future career path recommendations while ignoring current market trends. This work proposes a framework that combines reinforcement learning, Bayesian inference, and gradient boosting to provide recommendations on how contractors *should* behave when choosing new skills to learn. Compared with standard recommender systems, this framework does not learn from previous (often poor) behaviors to make future recommendations. Instead, it relies on a Markov decision process to operate on a graph of feasible actions and dynamically recommend profitable career paths. The framework uses market information to identify current trends and project future wages. Based on this information, it recommends feasible, relevant actions that a contractor can take to learn new, in-demand skills. Evaluation of the framework on 1.73 million job applications from an online labor market shows that its implementation could increase (1) the marketplace's revenue by up to 6%, (2) contractors' wages by 22%, and (3) the diversity of new skill acquisitions by 47%. A comparison with alternative recommender systems highlights the limitations of approaches that make recommendations based on previously observed learning behaviors.

History: Accepted by Chris Forman, information systems.

Supplemental Material: The online appendices are available at <https://doi.org/10.1287/mnsc.2020.3727>.

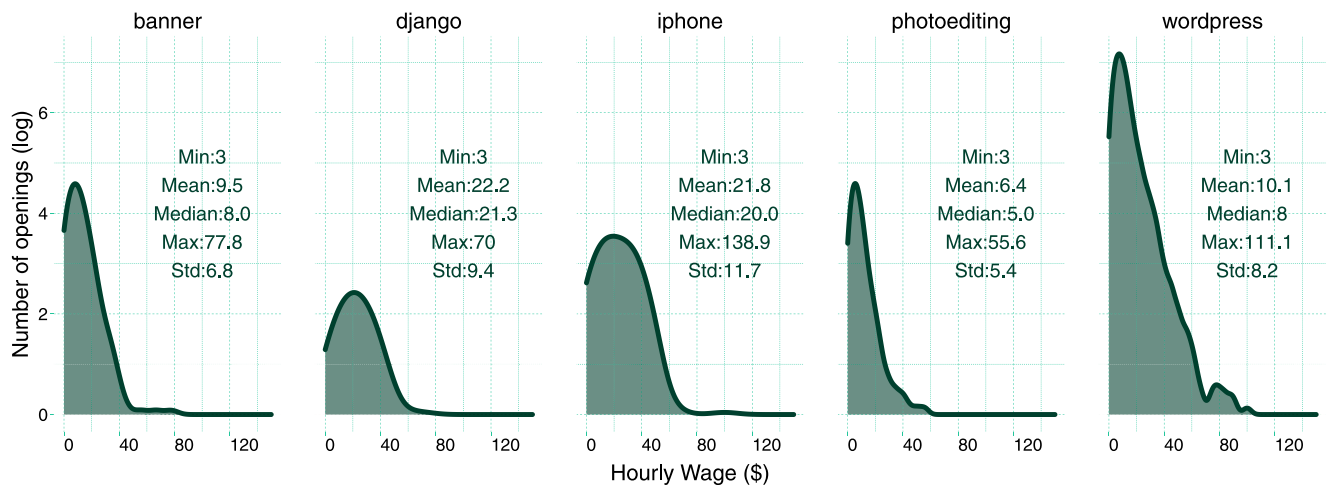
Keywords: skill recommendations • artificial intelligence • online labor markets • Markov decision process • information systems: enabling technologies

1. Introduction

Online labor markets such as Upwork and PeoplePerHour facilitate independent employers to connect with contractors around the world to accomplish a diverse set of tasks. These tasks span multiple categories, including data science, technical writing, digital marketing, and web development. On par with other online platforms, online labor markets grew exponentially during the past decade (Upwork 2014, Freelancers Union and Upwork 2017). Upwork, for example, has 14 million contractors and 5 million employers. These contractors complete more than three million tasks annually—a total transaction volume of \$1 billion (Lauren 2017, Brier and Pearson 2018). This growth will likely continue (if not accelerate) in the future, as automation and the sharing economy structure the future of work (Sundararajan 2016, Jain et al. 2018).

In these digital workplaces, a skill's value (compensation wage) depends on market conditions, that is, the skill's demand and supply. Market supply and demand vary widely, both *across* and *within* skills. This variation creates heterogeneous distributions of compensation wages: Figure 1 presents the hourly compensation wages for five representative skills in the focal data set. The *y*-axis shows the number of openings (in log scale), which is a rough proxy of market demand. The *x*-axis presents the hourly wage in U.S. dollars paid to the *hired* contractors of these openings. Expectedly, there is a high variance across skills. The median wage for photo editing is \$5, whereas the median wage for iphone is \$20; the total number of openings for wordpress is 19,828, whereas for django is 339. Similarly, there is a significant variance in the *shape* of the wage distributions within each skill: From highly concentrated (photo editing)

Figure 1. (Color online) Heterogeneous Compensation Wages of Skills in Online Labor Markets



Note. Variability exists both *within* and *across* skills.

to lognormal-distributed (iphone, banner, django), to gamma-distributed (wordpress) wages. This within-skill variance captures the global dimension of the market (i.e., contractors from different countries charge different premiums), as well as the existence of contractors with different levels of expertise.

The dynamic nature of online labor markets further amplifies this observed heterogeneity. To keep up with shifting market conditions (i.e., demand and supply for new and old skills), contractors need to continuously keep reskilling themselves (Autor et al. 1998, Autor 2001, Kuhn and Skuterud 2004, Stevenson 2009, Oliver 2015). During this reskilling process, contractors face a crucial question: *What skills should they acquire to remain (or become) competitive in such a dynamic, heterogeneous environment?*

Observational data suggest that many contractors often make poor decisions when they choose to learn new skills. Figure 2 shows the distribution of the

average wage improvement after the acquisition of a new set of skills. On average, new skill acquisition yields a small wage increase (4.3%). However, a significant fraction of contractors experiences a wage decrease. A variety of factors could explain this observation, including information asymmetry about the value of newly acquired skills, search costs associated with the decision to acquire new skills, and the contractors' inability to (1) predict future demand and (2) assess their competence in learning individual skills.

Recommendation frameworks can address these issues by providing curated sets of relevant and profitable learning choices. Perhaps the most straightforward skill recommendation approach (often adopted by online labor markets) is to advertise skill-specific wages (Upwork Enterprise 2017). Such reports implicitly guide contractors to acquire skills that are in high demand. This approach has two significant

Figure 2. (Color online) Wage Improvements from New Skill Acquisitions



Notes. Contractors choose to acquire new skills that lead, on average, to a 4.3% increase in wages. Often, such acquisitions decrease the contractor's market value (decrease in compensation wages).

shortcomings: First, not every skill fits every contractor. As an extreme example, recommending machine learning to a contractor with data-entry expertise would not be very useful, as it is practically infeasible for a contractor without a relevant background to learn machine learning. Second, such skill reports are not diverse, as they often present a small set of highly valued skills (Upwork Enterprise 2017). As a result, they guide contractors to follow similar career paths that could saturate the market with similar contractors.

Conventional recommender systems (e.g., collaborative filtering) also have shortcomings when applied in the focal context. Because they rely on observed data to suggest “*skills that similar contractors have learned in the past*,” they limit the emergence of skill recommendations that are promising and have increased demand. Furthermore, because these systems model prior user behavior, their recommendations often reflect poor choices that contractors make when acquiring new skills (Figure 2). Besides, similar to skill-specific wage reports, these systems often result in nondiverse recommendations of popular “items” (Fleder and Hosanagar 2009, Adamopoulos 2014, Adamopoulos and Tuzhilin 2014, Adamopoulos et al. 2015). Hence, they could also saturate the market with contractors with similar skillsets.

This work proposes a framework that resolves these shortcomings by providing *model-driven* recommendations on how contractors *should* behave. Compared with standard recommender systems, this framework does not learn from contractors’ previous (often poor) behaviors to make future recommendations. Instead, it acts as a career development adviser that attempts to *maximize the future expected earnings* of contractors by providing *feasible* and *relevant* career paths. The framework uses market information to estimate current and future wages for different skills, while it uses observed skillsets to identify the feasible actions that a contractor can take to learn new skills.

The proposed approach relies on a Markov decision process (MDP) to dynamically recommend actionable career paths. At each period, the MDP provides recommendations on how contractors *should* spend their next period: (1) learn a new set of skills, (2) work using their current set of skills, or (3) do both and learn while working. The current state of each contractor (i.e., contractor’s skillset) defines the set of *feasible actions*, each of which results in an *immediate reward* and a stochastic transition to a new state. Feasible new states represent *observed supersets* of the contractor’s current state. Individual contractor characteristics define the transition probabilities to these states. Overall, the framework identifies actionable career development paths by combining techniques from

reinforcement learning, Bayesian inference, word embeddings, and gradient boosting.

To evaluate the performance of the framework, we use a transactional data set of 1.73 million job applications from a leading online labor market. Based on this data set, we devise and run an extensive set of simulations that control for future fluctuations of supply, demand, and wages due to the recommendations of the MDP framework. The results show that our approach can increase the revenue of the marketplace by 1.75%–6% through the increased availability of contractors with in-demand, scarce skills. Besides, contractors who follow the recommended career paths can expect an average increase of 22% in their compensation wages. Due to the high stochasticity of the MDP framework and the fact that it controls for current (and future) trends of supply and demand, the resulting skill recommendations are, on average, 47% more diverse than the observed contractors’ skill acquisitions. Thereby, such recommendations are unlikely to saturate the marketplace with very similar skillsets. Finally, a comparison with six alternative recommender systems shows the advantages of the proposed approach by highlighting the limitations of systems that make recommendations based on previously observed learning behaviors.

This work contributes to the design research of digital workplaces and has various managerial implications. First, it demonstrates that a model-driven recommendation framework resolves multiple shortcomings of current “vanilla” machine learning (and other) approaches. This is only feasible because the framework does not model prior contractor behavior; instead, it uses marketplace domain knowledge along with estimates of contractors’ abilities to develop feasible, in-demand career path recommendations. Second, because the framework operates on an up-to-date evaluation of the current state of the marketplace, it dynamically adapts to situations in which individual skills lose or gain value. Such online adaptations increase the awareness of the contractors about the current and future state of the marketplace. Through better awareness, contractors can make better-informed decisions and increase their hireability in the marketplace. Similarly, due to a better supply of high-demand skills, employers can efficiently identify appropriate candidates, which, in turn, increases the transaction efficiency of the workplace.

2. Background

The recent growth of online labor markets (Agrawal et al. 2015) has motivated multidisciplinary research that focuses on a variety of problems. Examples include algorithms to manage contractors (Sheng et al. 2008, Horton and Chilton 2010, Ipeirotis et al. 2010, Mason and Watts 2010), running randomized

experiments at scale (Horton et al. 2011, Berinsky et al. 2012, Rand 2012), design choices that increase market transaction efficacy (Gopal et al. 2003, Arora and Forman 2007, Dey et al. 2010, Horton 2010, Allon et al. 2012, Chen and Horton 2016, Liang et al. 2016, Kokkodis 2020), and drivers of hiring decisions (Banker and Hwang 2008, Yoganarasimhan 2013, Moreno and Terwiesch 2014, Lin et al. 2016). Despite these efforts, little research has investigated the development of human capital through skill acquisitions in these environments.

2.1. Investment in Human Capital

Investing in human capital drives economic growth (Schultz 1961, Becker 1962, Blundell et al. 1999). For societies, public investment in human capital reduces income inequality (Glomm and Ravikumar 1992). For organizations, such investments lead to competitive advantage (Lepak and Snell 1999). For individuals, it allows for changes in career paths that often lead to higher earnings (Johnson 1970, Stevenson 2009, Agarwal and Ohyama 2013), while it also enhances entrepreneurial performance (Bosma et al. 2004). A skill's value, however, is context-dependent, as organizations weight combinations of skills differently and according to their needs (Lazear 2009).

Individual investment in human capital through new skill acquisition is particularly relevant in digital workplaces, where new skills are born and old skills die faster than ever (Autor et al. 1998, Autor 2001, Oliver 2015, Kokkodis and Ipeirotis 2016). To remain marketable, contractors must be diligently and continuously reeducating and reskilling themselves to keep up with the shifting labor market needs (Kuhn and Skuterud 2004, Stevenson 2009, Oliver 2015, Kokkodis 2020).

But choosing skills to learn is an inherently hard task (Gati and Levin 2014). First, contractors need to assess their ability to learn specific skills. Such assessments are bound to interpretations of past experiences (Gati and Levin 2014) and to the contractors' current levels of confidence that structure their beliefs about their learning abilities (Fischhoff et al. 1977, Cooper et al. 1988, Odean 1998, Johnson and Fowler 2011, Myers 2011). Second, contractors need to (1) assess current and future market conditions and (2) choose skills that would justify their investment in time (and possibly) money (Rosen 1983, Becker 1994). Because contractors cannot have a complete picture of the marketplace, assessment of the market value of skills happens under conditions of *information asymmetry* (Akerlof 1970). Such choices under uncertainty lead to poor outcomes (Wiltbank et al. 2009). Third, choosing skills to learn induces a *search cost* (Smith et al. 1999), which is exemplified in digital workplaces as these markets are highly

heterogeneous in terms of skills and qualifications (Kokkodis and Ipeirotis 2014, 2016).

2.2. Current Recommender Systems

Recommender systems reduce *search costs* and often resolve *information asymmetries* by providing accurate and efficient recommendations (Fleder and Hosanagar 2009, Pathak et al. 2010, Brynjolfsson et al. 2011). State-of-the-art algorithms in online platforms (e.g., ecommerce, streaming services, social networks, online labor markets) provide a curated set of product or service recommendations—reduced search cost—that the targeted user will likely enjoy—reduced information asymmetry (Schafer et al. 1999; Bennett et al. 2007; Phelan et al. 2009; Bian and Holtzman 2011; Kokkodis and Ipeirotis 2013; Adamopoulos and Tuzhilin 2015a; Kokkodis et al. 2015; Gomez-Urbe and Hunt 2016; Kokkodis 2018, 2019; Kokkodis and Ipeirotis 2020). Current recommender systems divide into three categories: content-based, collaborative, and hybrid (Adomavicius and Tuzhilin 2005, Kantor et al. 2011).

Content-based systems recommend items that are similar to items that the targeted user has liked in the past (Adomavicius and Tuzhilin 2005). Examples include the recommendation of books (Mooney and Roy 2000), web pages (Pazzani and Billsus 1997), and news (Lang 1995). These systems experience the cold-start problem (Adomavicius and Tuzhilin 2005, Lam et al. 2008), where recommendations for new users with no (or very limited) history are infeasible. Furthermore, such systems tend to overspecialize recommendations, as they only recommend items that are similar to those previously rated highly by the user (Adomavicius and Tuzhilin 2005; Adamopoulos 2013a, 2014; Adamopoulos and Tuzhilin 2014, 2015b).

Collaborative recommender systems suggest items that similar users with the targeted user have liked in the past (Billsus and Pazzani 1998, Breese et al. 1998). Such approaches split into two general classes: memory-based and model-based. Memory-based algorithms “make rating predictions based on the entire collection of previously rated items by the users” (Adomavicius and Tuzhilin 2005, p. 738; Breese et al. 1998, Delgado and Ishii 1999). Model-based approaches use the collection of ratings to learn a model, which then makes rating predictions (Billsus and Pazzani 1998, Breese et al. 1998, Getoor et al. 1999). Such systems have been very successful in ecommerce and other digital platforms (Schafer et al. 1999, Gomez-Urbe and Hunt 2016, Mustafa et al. 2017). Similarly to content-based systems, collaborative approaches experience the cold-start problem, for both new users and new items (Adomavicius and Tuzhilin 2005).

Hybrid systems combine notions from content-based with characteristics of collaborative approaches to provide systems that overcome some of

the limitations of the two approaches (Balabanović and Shoham 1997, Si and Jin 2003, Adomavicius and Tuzhilin 2005, Tso-Sutter et al. 2008, Sahoo et al. 2012). For instance, hybrid systems can combine outputs of collaborative and content-based systems through linear weighting or other voting schemes (Claypool et al. 1999, Pazzani 1999). Overall, hybrid systems outperform pure collaborative and content-based methods and tend to provide more accurate recommendations (Balabanović and Shoham 1997, Pazzani 1999, Melville et al. 2002).

2.3. Recommending Skills

These recommender systems rely on previously observed user behavior to recommend actions for future user behavior. For instance, they learn from previous products that users liked to recommend future products. Acquiring new skills and developing human capital, however, are fundamentally different from buying products or services. This contextual uniqueness of our problem renders current recommender systems not directly applicable. Direct application of these systems would result in recommendations such as “*learn skills that similar contractors have learned in the past.*” Such recommendations would limit the emergence of new skills that are promising and have increased demand. Furthermore, because contractors often make poor choices (Figure 2, Section 2.1), learned skill-acquisition patterns would incorporate these observed poor decisions. At the same time, vanilla recommender systems would likely yield nondiverse recommendations of a few popular skills (Fleder and Hosanagar 2009).

Besides, context-specific peculiarities further reduce the applicability of conventional recommender systems. Such recommender frameworks predict the “rating” a “user” would assign to an “item” (Ricci et al. 2011). To apply these systems in our context, we need to map contractors to “users,” skillsets to “items,” and contractors’ compensations to “ratings.” This application would result in recommending similar skills to similar users, ignoring individual characteristics such as the contractor’s ability to learn new skills. Besides, it would assume that the interactions between the “users” and the “items” happen only once (i.e., a user watches a given movie or buys a given product only once), whereas the majority of the contractors use the same skills over and over again. Finally, such a specification would ignore actions that do not involve the learning of a new skill, such as working or learning while working.

The proposed recommendation framework overcomes these shortcomings and proposes an approach that fits the context-specific peculiarities of new skill acquisition. Specifically, it does not rely on previous, often poor behavior to recommend future behavior.

Instead, it uses information from the market to identify feasible and likely profitable future career paths. Because it is built particularly for skill acquisitions, it allows for the sequential evolution of contractors through learning, working, and learning while working. Since our approach leverages observed information to estimate a contractor’s learning ability and likelihood of successful skill acquisition, it results in stochastic recommendations that are significantly different between seemingly similar contractors. The result is a diversified career development robo-advisor that models how contractors *should* behave in terms of skill acquisition.

3. Research Context

The robo-advisor operates on a space of observed skillsets. To showcase, we use a transnational data set from an online labor market.

3.1. Data Set

The focal snapshot includes 1,727,680 job applications that led to more than 95,588 completed tasks. It covers transactions over four years, and it includes complete information in terms of job postings, job applications, hires, contracts, and evaluations. The analyzed tasks span all available task categories, including software development, web development, logo design, technical writing, and data analytics. The total number of observed skills is 215, which generate a total of $|\mathcal{S}| = 3,386$ observed skillsets. Contractors come from 182 countries and earn, on average, \$13.6 per hour. Table 1 summarizes this data set.

3.2. Model-Free Evidence

The process of learning new skills is unobserved to the market. Contractors often signal a new skill acquisition by updating their profile to list a new skill. Figure 3(a) shows that, over the four years of the focal data set, contractors listed on average 5.6 new skills on their profiles—87% of the contractors listed at least one or more new skills. Given that online profiles frequently contain inaccurate information (Toma et al. 2008), contractors might list skills that they do not possess. Figure 3(b) shows the percentage of newly listed skills that contractors have subsequently worked on: 83% of the contractors who listed new skills on their profiles worked on at least one task that required one of those new skills.

Of course, there is a nontrivial chance that contractors indeed learn a new skill, but for various reasons (e.g., demand for the new skill), they do not work on that skill. Figure 3(c) shows how long it takes for a contractor to exercise a newly listed skill. On average, contractors work on newly acquired skills 43 days after the profile listing (median 28 days). Within the first six months from the observable skill

Table 1. Data Overview

	Observations	Minimum	Mean	Median	Maximum	Standard deviation
Skills per user	1,727,680	1	9.6	9	61	5.9
Required skills per opening	135,354	1	2.4	2	21	1.7
Hourly wage	95,588	3	13.6	11.1	219	10.6
Total applications	1,727,680					
Total openings	135,354					
Completed tasks	95,588					
Unique skills (G)	215					
Unique observed skillsets (S)	3,386					

Notes. Contractors come from 182 countries. The data set spans four years.

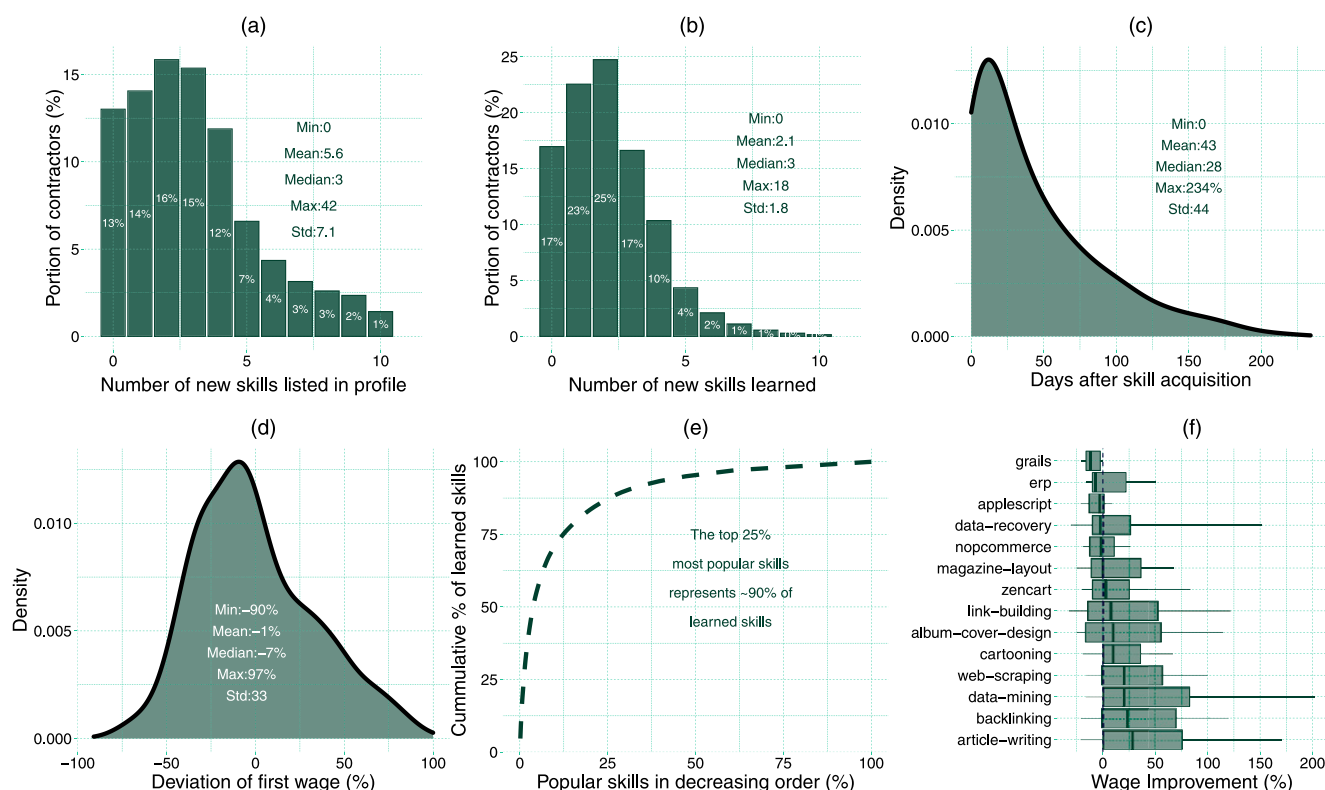
acquisition, 99% of the contractors work on such skills. Furthermore, Figure 3(d) shows that overall, when contractors acquire new skills, they start working on them with slightly lower-than-average premiums (−1%).

Currently, online labor markets do not offer career path recommendations. Instead, they usually promote individual skills through various press releases (Upwork Enterprise 2017, Schultz 2018). This can potentially

result in the saturation of popular skills. Figure 3(e) shows that the top 25% most popular skills represent 90% of the observed learning choices; the top five most popular skills constitute 25% of the learning decisions.

Contractors often do not have access to the current (and future) market value of skills. When choosing to learn new skills, contractors do it myopically, with noisy information that they can gather about the

Figure 3. (Color online) Model-Free Evidence of Observed New Skill Acquisition



Notes. Panel (a) shows the distribution of total number of new skills listed on contractors' profiles. Within this newly set of skills, panel (b) shows how many skills contractors actually used in the market ("Number of new skills learned"). Panel (c) shows the distribution of how many days it takes for a contractor to work on a new skill. Panel (d) shows the wage deviation of the first completed job that requires the newly acquired skill from the average skill's wage. Panel (e) shows that a few skills constitute the majority of learning choices, which can potentially result in skill saturation. Panel (f) shows that contractors often make poor choices that result in wage decreases.

skills' market value. Figure 3(f) shows evidence of the existence of such information asymmetry. The x -axis shows the average wage improvement after the acquisition of a new skill. The y -axis lists a randomly selected set of skills. The variance both within and across these skills reflects the uncertainty of outcomes when contractors choose to learn new skills.

To summarize, data evidence shows that contractors are dynamic entities that invest in learning new skills. Their current behavior, however, results in poor learning choices that can potentially (1) saturate popular skills and (2) yield negative returns of investment. This model-free evidence further highlights the need for a recommendation framework that uses market information to reduce uncertainty in learning decisions and guide contractors to choose skills that will not become saturated and will likely yield positive returns.

4. A Career Path Adviser

The proposed career path adviser models how contractors *should* behave when acquiring new skills. At any moment in time t , the framework provides recommendations on how contractors should spend their next period. They can work using their existing set of skills \mathbf{S} , learn a new set of skills \mathbf{N} , or combine working and learning, by working for a λ fraction of their time and learning for the rest $1 - \lambda$. Contractors' decisions lead *stochastically* to different outcomes. For example, a contractor who chooses to learn new skills \mathbf{N} and succeeds, ends up with a new skillset $\mathbf{S}' = \mathbf{S} \cup \mathbf{N}$. This skillset \mathbf{S}' may generate a higher payoff for future tasks, in exchange for "lost" time spent learning during the acquisition of the new skills. Similarly, a contractor who chooses to learn a new skill while working in parallel might fail in learning and end up not monetizing the studying time.

We model this stochasticity of the contractors' actions in terms of payoffs and outcomes as a MDP. To formally define an MDP, we need:

- *A set of states \mathbb{S}* : Each state is defined by the skillset of a contractor \mathbf{S} . Set \mathbb{S} is defined by the observed skillsets in the marketplace (see Table 1).
- *A set of feasible actions $A(\mathbf{S})$ for every state \mathbf{S}* : An action $\alpha := (\lambda, \mathbf{N})$, $\alpha \in A(\mathbf{S})$, corresponds to learning a set of new skills \mathbf{N} while allocating λ percent of the time to working.
- *Transition probabilities between states $\Pr(\mathbf{S}, \mathbf{S}'|\alpha)$, $\mathbf{S}' = \mathbf{S} \cup \mathbf{N}$, $\alpha \in A(\mathbf{S})$* : Transition probabilities vary according to (1) the learning affinity of the new skills \mathbf{N} to the existing skillset \mathbf{S} ($L(\mathbf{N}|\mathbf{S})$), (2) the unobserved learning ability of the contractor (d), (3) the fraction of time spent on working (λ), and (4) the country of the contractor (c).

- *A reward function $R(\mathbf{S}, \alpha)$, $\forall \alpha \in A(\mathbf{S})$* : Immediate rewards depend on the fraction of time λ that a contractor chooses to work while in state \mathbf{S} .

Table 2 summarizes the notation. Figure 4 shows the components of the MDP. The observed skillsets in the marketplace directly define the space of available states \mathbb{S} . Sections 4.1, 4.2, and 4.3 present the details for structuring (1) the set of *feasible* actions $A(\mathbf{S})$ for each state \mathbf{S} , (2) the transition probabilities $\Pr(\mathbf{S}, \mathbf{S}'|\alpha)$, and (3) the reward function of each action $R(\mathbf{S}, \alpha)$. Section 4.4 presents algorithms that efficiently solve the MDP to recommend personalized paths (i.e., policies) that maximize the long-term expected rewards of each contractor in the marketplace. Finally, Section 4.5 presents illustrative examples, and Section 4.6 draws an analogy between the proposed robo-advisor and navigation systems to summarize the robo-advisor's functionality and benefits.

4.1. Component 1: Feasible Actions

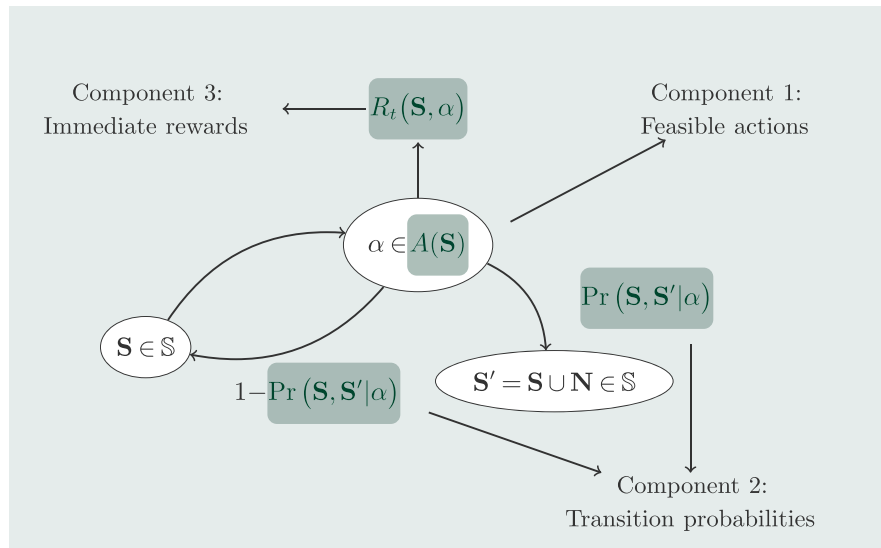
In an online labor market, the number of available skills is large, often in the hundreds or even thousands. However, these skills are not uniformly distributed across skillsets. Instead, clusters of skills are usually found together, whereas combinations of clusters create disjoint observed skillsets.

Data Evidence 1. Let us assume that a universe of skills is the following: $\mathbf{G} = \{\text{wordpress, marketing, blog, twitter, logo, management, translating, sql, asp, seo}\}$. The potential state space consists of $|\mathbb{S}^p| = 2^{|\mathbf{G}|} = 2^{10} = 1,024$ possible states. However, in the focal data set we only observe $|\mathbb{S}| = 68$ states. This is a direct consequence of the fact that contractors usually specialize in skills that are correlated. For instance, in the focal data set we never observe contractors

Table 2. Notation Summary

Symbol	Description
\mathbf{G}	Universe of available skills
\mathbb{S}	The set of available states (observed skillsets)
\mathbf{S}	The set of contractor skills, effectively a state of the MDP
\mathbf{N}	A set of new skills that a contractor learns
\mathbf{S}'	$\mathbf{S} \cup \mathbf{N}$
λ	Fraction of time invested in working
$A(\mathbf{S})$	The set of feasible actions from state \mathbf{S}
$\alpha := (\lambda, \mathbf{N})$	The action of learning new skills \mathbf{N} while working λ percent of the time
$R_t(\mathbf{S}, \alpha)$	Rewards of action α at state \mathbf{S} and time t
$\Pr(\mathbf{S}, \mathbf{S}' \alpha)$	The probability of transitioning from state \mathbf{S} to state \mathbf{S}' given action α
$L(\mathbf{N} \mathbf{S})$	The learning affinity of the new skillset \mathbf{N} to state \mathbf{S}
d	The unobserved individual ability of a contractor to learn new skills
c	The country of a contractor
i	Contractor index

Figure 4. (Color online) The Three Components of the MDP



Notes. The first component models the set of feasible actions $A(S)$ from each state S . The second component defines the transition probabilities of each action. The third component of the MDP estimates the immediate rewards of each action. Assume that S is the current state of a contractor. The contractor takes an action α from the set of feasible actions and stochastically transitions to either a new state S' or remains in state S . The contractor's action emits an immediate reward. The ultimate goal of the MDP framework is to identify a set of actions (i.e., a personalized policy) that yields high long-term expected rewards for any given contractor.

who complete tasks that require $S = \{\text{asp, blog}\}$. In fact, this means that we never observe contractors who complete tasks that require $S = \{\text{asp, blog, N}\}$, where $N \in \mathcal{P}(G - S)$ and \mathcal{P} is the subset's power set. The gap between the possible number of states and the observed ones becomes greater as the number of available skills ($|G|$) increases.

When contractors from state S choose to acquire new skills N , they stochastically transition to a new state $S' = S \cup N$. Hence, plausible transitions require the new state S' to be a superstate of the current state S . This additional constraint further limits the set of feasible transitions from each state $S \in \mathcal{S}$.

Data Evidence 2. Figure 5 shows the six feasible transitions from state $S = \{\text{sketching}\}$. Considering only the five skills that appear in this figure, $G = \{\text{sketching, drawing, illustration, cartooning, logo}\}$, the number of possible superstates that include skillset S is $2^4 - 1 = 15$. However,

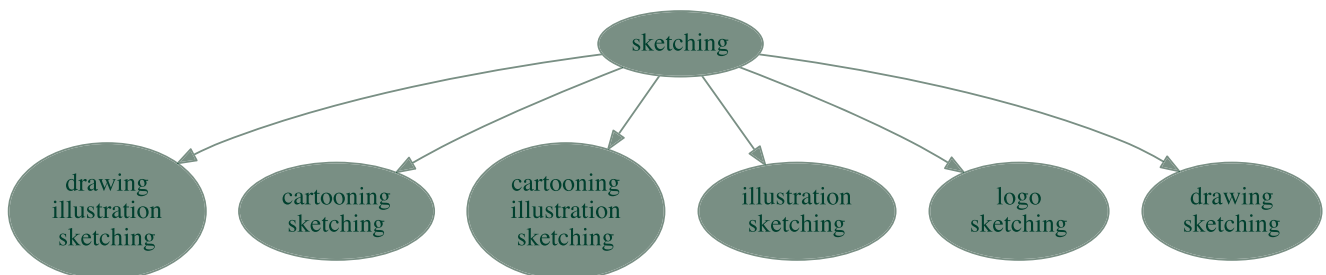
in the focal data set we only observe 6 out of these 15 supersets. Hence, learning actions from state $S = \{\text{sketching}\}$ can only lead to one of these six observed states of Figure 5. In general, in the extensive screenshot of the marketplace we analyze in this study there are 44,190 feasible transitions as a result of the $G = 215$ skills and the $\mathcal{S} = 3,386$ observed states (Table 1). Figure 6 shows that feasible transitions per state follow a power law distribution. The median of this distribution is 7 (average is 19): 50% of the available states in the focal data set have seven or more feasible transitions.

Formally, the set of feasible actions from state S is

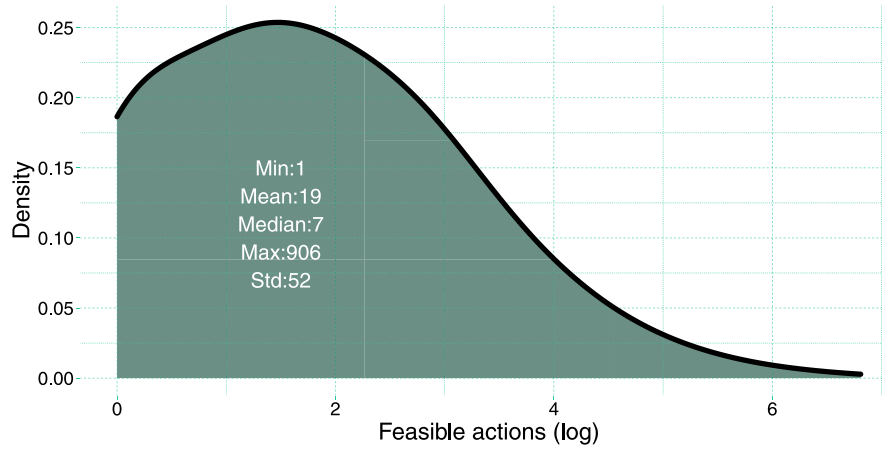
$$A(S) = \{\alpha : S \cup N \in \mathcal{S}\}, \quad \alpha := (\lambda, N). \quad (1)$$

This set of actions results in feasible transitions that include only *observed* (i.e., $S \cup N \in \mathcal{S}$) superstates of the current state S . On the one hand, there is a nonzero probability that this feasible set of actions might miss

Figure 5. (Color online) Feasible Transitions from State {sketching}



Notes. There are only six feasible transitions from an initial state $S = \{\text{sketching}\}$. Omitted combinations of skills are not observed in the data.

Figure 6. (Color online) Number of Feasible Transitions per State

Note. The number of feasible transitions per state follows a power law distribution.

potentially lucrative combinations of skills that the market does not observe. On the other hand, constraining transitions to observed states (1) is necessary for a data-driven representation of the expected rewards (Section 4.3) and (2) results in the tractable estimation of the framework parameters (Section 4.4).

4.2. Component 2: Transition Probabilities

The MDP allows contractors who take feasible actions to stochastically transition to new states. For a given contractor in state \mathbf{S} who takes an action α the transition probability to move to state $\mathbf{S}' = \mathbf{S} \cup \mathbf{N}$ depends on (1) the learning affinity $L(\mathbf{N}|\mathbf{S})$ of the new skills \mathbf{N} to the contractor's current state \mathbf{S} , (2) the contractor's unobserved learning ability d , (3) the fraction of time λ that the contractor spends on working while learning, and (4) the country of the contractor (c). A logistic regression model estimates this probability as follows:

$$\Pr(\mathbf{S}, \mathbf{S}'|\alpha) = \frac{\exp(\beta_0 + \beta_1 L(\mathbf{N}|\mathbf{S}) + \beta_2 d + \beta_3 \lambda + \beta_4 c)}{1 + \exp(\beta_0 + \beta_1 L(\mathbf{N}|\mathbf{S}) + \beta_2 d + \beta_3 \lambda + \beta_4 c)}. \quad (2)$$

In the previous equation λ is given by the contractor's action and c is observed. In the next paragraphs, we define and estimate the learning affinity of a new state $L(\mathbf{N}|\mathbf{S})$ and the unobserved learning ability of a contractor d .

4.2.1. Learning Affinity. Skills are contextually connected. For instance, knowing how to program in java makes learning C# (which is a very similar programming language) an easy task. On the contrary, learning cartooning requires an entirely different set of abilities than programming; thereby, it is not contextually similar to java. To uncover such contextual connections between different sets of skills,

we draw on the field of natural language processing, and we use a distributed representation of words model (word embeddings (W2V); Mikolov et al. 2013). (Appendix A presents alternative modeling approaches.) W2V embeds words from a vocabulary into a lower-dimensional space, in which semantically similar words appear close to each other, whereas semantically dissimilar words appear far away from each other (Mikolov et al. 2013). In the context of this work, a "skill" maps to a "word" and a "skillset" to a "document." As a result, the analyzed corpus consists of contractor skillsets. Based on this representation, W2V projects contextually similar skills close to each other in a low-dimensional vector space of real numbers.

To create skillset representations, we simply aggregate the vectors of the mapped skills (Appendix A, Equation (13)). Comparison between skillsets happens by computing the cosine similarity of their corresponding vectors. Formally, we define the *learning affinity* $L(\mathbf{N}|\mathbf{S})$ of a new set of skills \mathbf{N} to the existing set of skills \mathbf{S} as

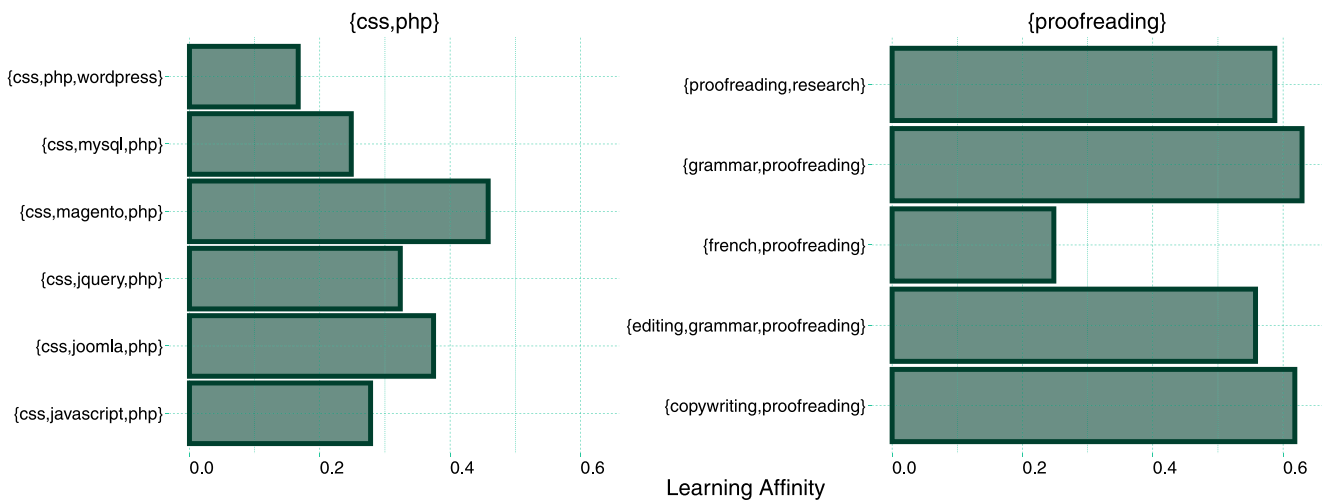
$$L(\mathbf{N}|\mathbf{S}) = \cos(v(\mathbf{N}), v(\mathbf{S})), \quad (3)$$

where $v(\cdot)$ is the vector representation of the skillset and $\cos(\cdot)$ is the cosine similarity.

Data Evidence 3. Figure 7 shows the computed learning affinities for a randomly chosen subset of feasible transitions from states $\mathbf{S}_1 = \{\text{css, php}\}$ and $\mathbf{S}_2 = \{\text{proofreading}\}$. This affinity ranges between 0.14 and 0.46 for \mathbf{S}_1 and between 0.25 and 0.63 for \mathbf{S}_2 . On average, in the focal data set feasible transition states have a learning affinity of 0.31 (minimum −0.34, maximum 0.69, and median 0.31).

4.2.2. Unobserved Learning Ability (d). Every contractor has a different unobserved (latent) ability to learn new skills. The focal data set allows the

Figure 7. (Color online) Learning Affinities for Skillsets {css, php} and {proofreading}



Notes. The figures show a subset of the feasible transition states from skillsets {css, php} and {proofreading}. Learning affinity in the two figures ranges between 0.14 and 0.63.

observation of contractors' histories, where we often find evidence of new skill acquisitions (Figure 3, (a) and (b)). Such skill acquisitions correlate with the latent learning ability of each contractor. Specifically, when a contractor acquires a new skill, we first observe the change on the contractor's profile. Sometimes, contractors also take certification tests, which they list as proof of expertise (Ipeirotis 2013). In the presence of listed certifications, we assume that a given contractor has successfully acquired the skills if the certification score of the contractor is in the upper 50%th percentile.¹ In the absence of certifications, we assume a successful skill acquisition if a contractor completes a job that requires the newly acquired set of skills and receives a high² feedback score. (Figure 3(c) shows that 99% of contractors work on their newly listed skills within the first six months. Hence, we only consider profile changes for which we have subsequent observations for more than six months.)

We model the sequence of observations for each contractor through a Bayesian inference framework. (Appendix B presents alternative modeling approaches.) In particular, we assume that the latent ability of a contractor to learn new skills (d) follows a prior distribution $d \sim \text{Beta}(\beta_1, \beta_2)$, $d \in (0, 1)$. For every new skill that contractors list on their profile, we get a Bernoulli trial that identifies whether the skill acquisition was successful. Based on the successful (κ) and unsuccessful ($n - \kappa$) skill acquisitions of the contractor, we estimate the posterior ability: $d \sim \text{Beta}(\beta_1 + \kappa, \beta_2 + n - \kappa)$.

To start this process, we use a nearest neighbor approach (on the listed skills and the contractor's country) to set the prior parameters β_1, β_2 to the average number of successes and failures of the most similar subgroup of contractors. Finally, every new

trial for skill acquisition updates the posterior estimate of the contractor's learning ability.

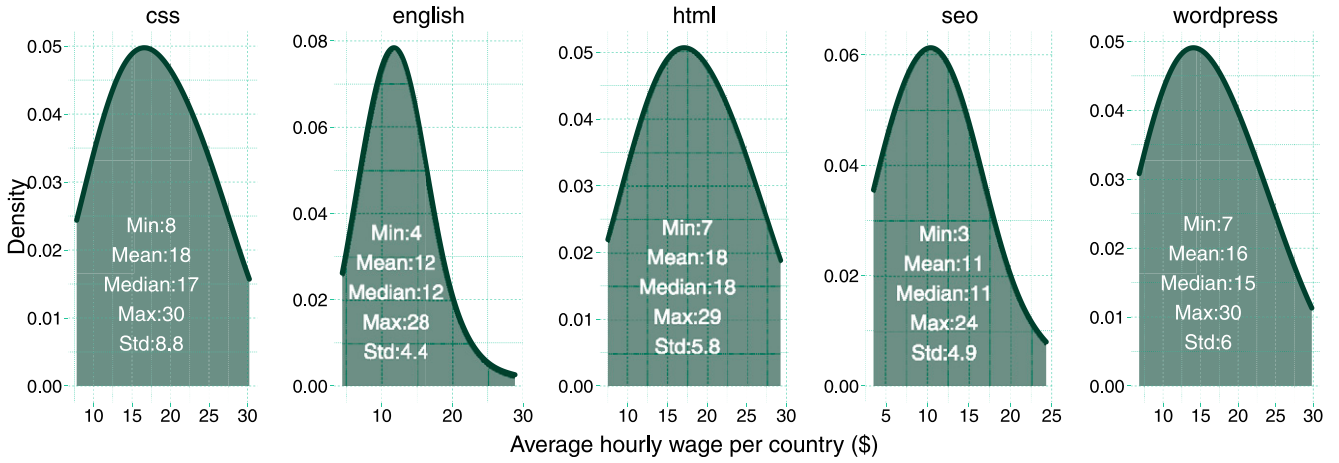
Once we have estimates for both the learning affinity of a feasible new state and the individual learning ability of a contractor, Equation (2) computes the transition probabilities. Appendix C presents the details.

4.3. Component 3: Immediate Rewards

The final component of the MDP is the function that defines the immediate rewards of each action. In an online labor market, contractors receive rewards by completing tasks and receiving compensation wages ($\lambda > 0$). On the contrary, they receive zero immediate rewards if they choose to learn a new skill without working ($\lambda = 0$).

Compensations are skillset- and location-dependent. For instance, all else being equal, a contractor that has a skillset in high demand is likely to charge more than a contractor with a skillset in lower demand (Marshall 2009). Similarly, all else being equal, a contractor from a developing country is likely to charge less than a contractor from the United States (Gefen and Carmel 2008).

Data Evidence 4. Consider contractors A and B. Contractor A completes a ruby-on-rails task; contractor B completes a linkbuilding task. In the focal data set, the average hourly wage of a ruby-on-rails task is \$22.68, whereas the average hourly wage of a linkbuilding task is \$5.04. In general, contractors' compensations are correlated with their set of skills. Similarly, compensations are bound to contractors' locations. Figure 8 shows the wage distributions for five different skills across the 182 countries in the data set. The average hourly wage per country ranges between \$5 and \$30.

Figure 8. (Color online) Average Hourly Wage per Country for Five Randomly Chosen Skills

Notes. The global dimension of the marketplace creates a high-variance, country-dependent distribution of wages. For these five skills, the average hourly wage per country ranges between \$5 and \$30.

These observations suggest that a proper compensation function needs to account for a contractor's (1) skillset (\mathbf{S}) and (2) country (c) as significant sources of heterogeneity. Specifically, for a contractor i at time t :

$$W_{it} = \bar{W}_t(\mathbf{S}, c) + \varepsilon_{it}, \quad (4)$$

where $\bar{W}_t(\mathbf{S}, c)$ is the average compensation wage for a contractor with skillset \mathbf{S} from country c , and ε_{it} is a zero-mean error term that captures individual characteristics of contractor i and market disturbances at time t .

Based on Equation (4), the expected immediate rewards for a contractor i who takes an action $\alpha \in A(\mathbf{S})$, $\alpha := (\lambda, N)$, at time t are

$$\begin{aligned} R_{it}(\mathbf{S}, \alpha) &= \mathbb{E}[\lambda \cdot W_{it}] \\ &= \lambda \cdot \mathbb{E}[\bar{W}_t(\mathbf{S}, c_i) + \varepsilon_{it}] \\ &= \lambda \cdot (\mathbb{E}[\bar{W}_t(\mathbf{S}, c_i)] + \mathbb{E}[\varepsilon_{it}]) \\ &= \lambda \cdot \bar{W}_t(\mathbf{S}, c_i). \end{aligned} \quad (5)$$

Conceptually, Equation (5) shows that the immediate rewards of a contractor i from country c_i who spends λ time on working are the expected revenue of the contractor during that period of time, $\lambda \cdot \bar{W}_t(\mathbf{S}, c_i)$.

We can decompose $\bar{W}_t(\mathbf{S}, c)$ into a time-dependent component of wage that depends on the skillset \mathbf{S} and a static component of wage that depends on the country c :

$$\bar{W}_t(\mathbf{S}, c) = \bar{W}_t(\mathbf{S}) + \bar{W}(c). \quad (6)$$

Skillset-specific compensation wages dynamically evolve, as they are bound to changing market conditions, that is, the supply and demand for each

skillset \mathbf{S} (Autor 2001, Oliver 2015). Hence, to identify career paths over a sequence of periods we need to predict future values for the supply, demand, and expected wages of each skillset. Specifically, at time t , we estimate the following:

$$\hat{\text{Sup}}_t(\mathbf{S}) = f_S \left(\sum_{j=1}^L \gamma_j \text{Sup}_{t-j}(\mathbf{S}) + \sum_{j=1}^L \beta_j \text{Dem}_{t-j}(\mathbf{S}) \right), \quad (7)$$

$$\hat{\text{Dem}}_t(\mathbf{S}) = f_D \left(\sum_{j=1}^L \gamma_j \text{Sup}_{t-j}(\mathbf{S}) + \sum_{j=1}^L \beta_j \text{Dem}_{t-j}(\mathbf{S}) \right), \quad (8)$$

$$\hat{W}_t(\mathbf{S}) = f_W(\hat{\text{Sup}}_t(\mathbf{S}), \hat{\text{Dem}}_t(\mathbf{S})). \quad (9)$$

Functions f_S, f_D, f_W capture an appropriate modeling choice for sequential observations (time series) and L is the number of lagged periods. Appendix D compares powerful time-series prediction algorithms, including ARIMAX, recurrent neural networks (LSTM), gradient boosting (XGBoost), and support vector machines regression (SVM Reg). Evaluation across different metrics shows that, for the focal data set, XGBoost works slightly better. Hence, in the rest of the paper, future wage, demand, and supply predictions occur through the trained XGBoost model. (Note that \hat{W} is an estimator of \bar{W} in Equation (6).)

4.4. Exploring Alternative Career Paths: Q-Learning

After estimating the three components of the MDP, we explore policies (career paths) that could maximize the future expected rewards of any given contractor. Formally, we search for a policy function $\pi^* : \mathbf{S} \mapsto A(\mathbf{S})$ that describes a path $\forall \mathbf{S} \in \mathbb{S}$ with respect to an action-value function Q . This procedure is called Q-learning (Watkins and Dayan 1992). We define an

action-value function $Q(S_t, \alpha_t)$ iteratively in the following way:

$$Q(S_t, \alpha_t) \leftarrow (1 - \delta)Q(S_t, \alpha_t) + \delta \cdot (R_t(S_t, \alpha_t) + \gamma \cdot \max_{\alpha} Q(S_{t+1}, \alpha)), \quad (10)$$

where δ is the learning rate that defines the rate at which new information overrides old information, γ is a discount factor for future rewards, and $Q(S_0, \alpha_0) = 0$. Conceptually, Equation (10) shows that the new long-term value is equal to the current reward $R_t(S_t, \alpha_t)$ plus all future rewards of the next state assuming that the contractor takes the best actions in the future (Hu 2016).³ (Appendix F discusses and compares alternative approaches for estimating an appropriate action-value function, including Monte Carlo, deep Q-learning, double Q-learning, and dueling Q-learning.)

Algorithm 1 (Exploring Alternative Career Paths: Q-Learning)

```

Initialize
 $Q(S, \alpha) := 0 \forall S \in \mathcal{S}, \alpha \in A(S), \delta = 0.1, \gamma = 0.99, \epsilon_0 = 0.9$ 
1: for each episode do
2:   for each contractor do
3:     Reset contractor to initial state
4:   end for
5:   for each step  $t$  of the current episode do
6:     for each contractor do
7:       Take  $\alpha_t$  action from state  $S_t$  according to
         the current value of  $\epsilon_t$  ( $\epsilon$ -greedy strategy)
         and function  $Q$ 
8:       Observe (predicted) rewards  $R_t(S_t, \alpha_t)$ 
9:       Observe new state  $S_{t+1}$  {Sampling from
         distribution  $\Pr(S_{t+1}|\alpha_t)$ }
10:       $Q(S_t, \alpha_t) \leftarrow (1 - \delta)Q(S_t, \alpha_t) + \delta \cdot (R_t(S_t, \alpha_t) + \gamma \cdot \max_{\alpha} Q(S_{t+1}, \alpha))$ 
11:      if  $S_{t+1} \neq S$  then
12:        Update supply for period  $t + 1$ 
13:      end if
14:    end for
15:     $\epsilon_{t+1} = \frac{\epsilon}{t+1}$  {Adaptive  $\epsilon$ -greedy}
16:    XGBoost predicts wage (rewards), supply,
      and demand for next period  $t + 1$ 
17:  end for
18: end for
19: return  $Q(S, \alpha) \forall S \in \mathcal{S}, \alpha \in A(S)$ 

```

Algorithm 1 describes the detailed procedure that explores alternative career paths and yields a state-action function Q . Equation (10) occurs in line 10; line 8 uses the reward function described in Section 4.3; line 9 uses the transition probabilities described in Section 4.2. In line 12, the algorithm updates the supply of contractors according to the observed transitions, while, in line 16, it uses the time-series predictive models (Section 4.3) to estimate future wages, demand, and supply values. Exploration of

alternative career paths happens in line 7 through an ϵ -greedy strategy: with probability ϵ , the algorithm chooses actions randomly from the available set of actions for state S_t , while, with probability $1 - \epsilon$, the algorithm chooses the action that maximizes future expected rewards according to the current Q -function. (Appendix E uses a grid-search approach to tune the learning hyperparameter δ , while it also compares alternative exploration strategies—adaptive ϵ -greedy versus Boltzmann exploration. Based on this analysis, for the focal implementation, we use an adaptive ϵ -greedy approach starting at $\epsilon = 0.9$, and we set the learning rate at $\delta = 0.1$. Finally, we set the discount factor at $\gamma = 0.99$.)

4.5. Illustrative Examples of Career Paths

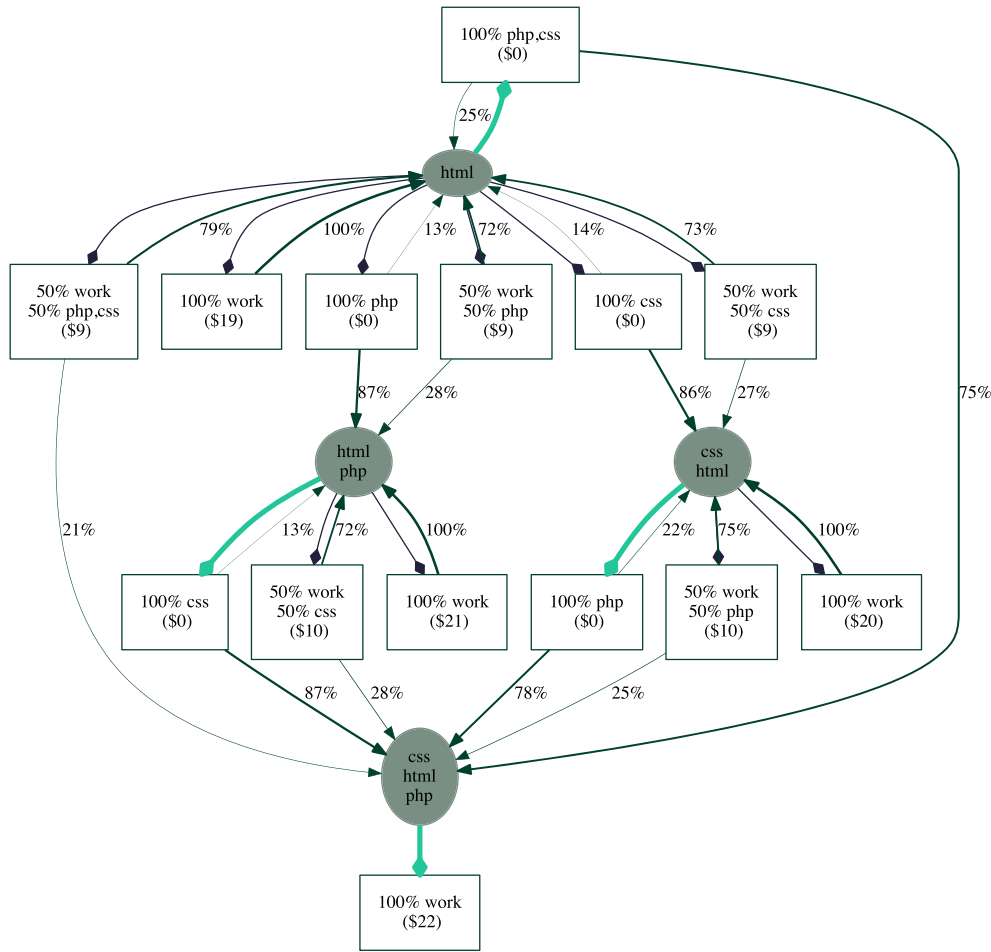
To visually demonstrate resulting career paths from the complete MDP framework, we use real data for two different fictional contractors from Greece. The first contractor has a strong ability to learn new skills ($d = 0.9$), whereas the second one has a lower learning ability ($d = 0.1$). For readability, this example considers only three randomly chosen skills: $\mathbf{G} = \{\text{css}, \text{html}, \text{php}\}$. The observed states are $\mathcal{S} = \{\{\text{html}\}, \{\text{html}, \text{php}\}, \{\text{css}, \text{html}\}, \{\text{css}, \text{html}, \text{php}\}\}$. At each state \mathbf{S} , a contractor chooses an action from the set $A(\mathbf{S})$. To provide a coherent visualization we consider $\lambda \in \{0, 0.5, 1\}$. (In practice the framework considers all values of $\lambda \in [0, 1]$.) We follow the process described earlier to estimate the transition probabilities and the expected immediate rewards. Algorithm 1 then derives the state-specific educated policies for the two contractors.

Figures 9 and 10 show the results for the high- and low-ability contractors, respectively. States are in shaded ellipses and feasible actions in rectangles. We pinpoint the available actions from each state through arrows with diamond endings, and we annotate the career paths with thick arrows with diamond endings. The printed wages are the average predicted wages according to Equation (9) across the following year for Greek contractors on each available state.

The two graphs differ in significant ways. First, the transition probabilities to new states are much higher for the first contractor who has high learning ability. This significant difference in probabilities results in entirely disjoint sets of educated actions. For the first contractor, the educated actions always include a learning component, intending to reach the most rewarding state $\mathbf{S} = \{\text{css}, \text{html}, \text{php}\}$. On the contrary, for the second contractor, the educated policy is to keep working, independent of the contractor's current state.

4.6. Framework Discussion

Before moving to the evaluation of the framework, it is essential to revisit its defining principles. First,

Figure 9. (Color online) Illustrative Career Paths for a Greek Contractor with Latent Ability $d = 0.9$ 

Notes. The universe of skills in this example is $G = \{\text{css}, \text{html}, \text{php}\}$. For this visualization we consider $\lambda \in \{0, 0.5, 1\}$. Arrows with diamond endings show the feasible actions from each state. Rectangles identify the immediate rewards from each action. Directed arrows represent the stochastic outcomes of each action. Thick arrows with diamond endings indicate the educated policy for this contractor (i.e., the recommended career path).

the framework does not model the current behavior of contractors. Instead, the framework builds a feasibility graph and advises contractors on how they *should* be learning new skills. The construction of the feasibility graph relies on observed skillsets in the marketplace. Based on this feasibility graph, the components of the MDP framework model the learning ability of contractors, the learning affinity of each new set of skills, and the future demand, supply, and wages that structure the long-term expected rewards of each action.

Conceptually, the MDP framework is similar to a navigation system. As in our context, in the absence of a navigation system, people often follow “suboptimal” routes to reach a destination (Lee and Cheng 2008). Hence, the navigation system does not rely on users’ previous behavior to recommend routes. Instead, it relies on current traffic information to identify on expectation the most efficient route to a destination. Similar to the MDP framework, navigation systems

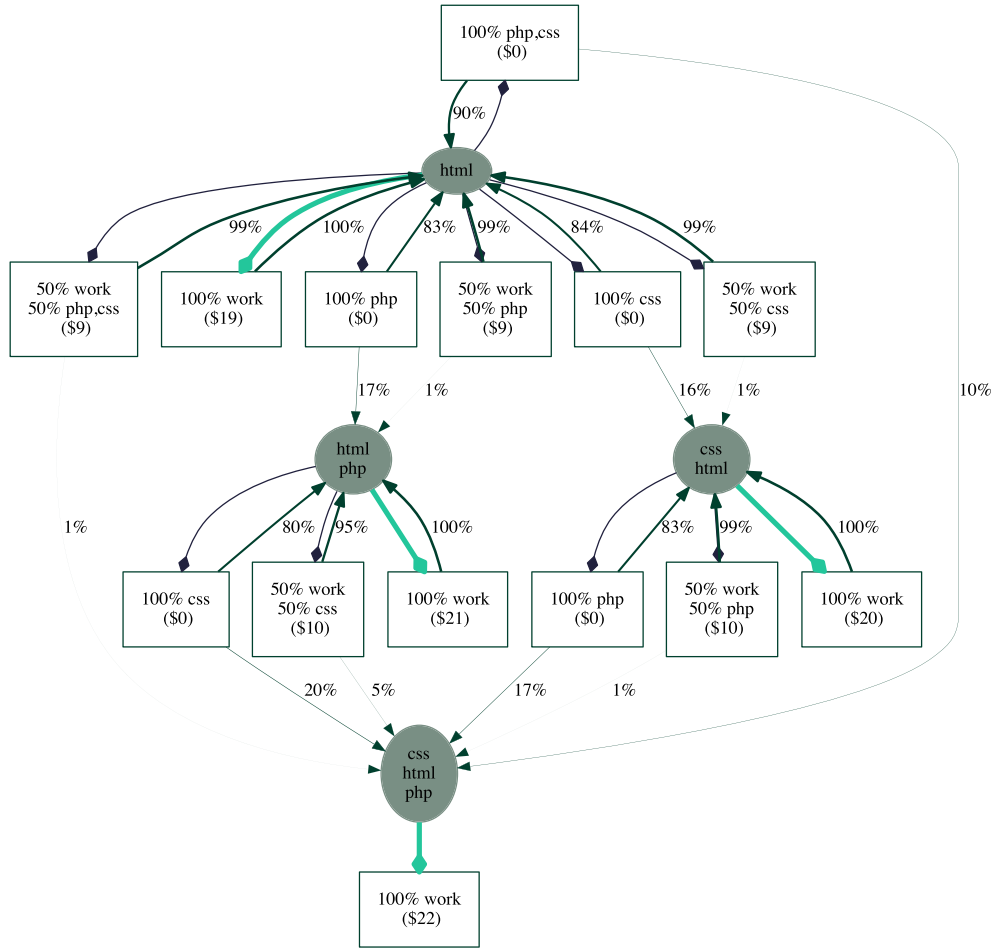
provide recommendations on how drivers *should* behave, ignoring (to a certain degree) the drivers’ previous behavioral patterns.

Although a navigation system does not rely on drivers to create the necessary maps, the MDP framework learns the feasibility graph from the observed skillsets in the marketplace. Hence, in theory, there may be states that are possible and lucrative but not yet observed in the marketplace (i.e., no contractor has the skills to form these states). In practice, however, given the scale of these markets (e.g., on Upwork, more than 14 million contractors), it is unlikely that such feasible lucrative states will remain hidden.

5. Performance Evaluation

To evaluate the complete framework we run a series of simulations over real and generated data and we compare its performance with various alternative recommender approaches.

Figure 10. (Color online) Illustrative Career Paths for a Greek Contractor with Latent Ability $d = 0.1$



Notes. The universe of skills in this example is $G = \{\text{css}, \text{html}, \text{php}\}$. For this visualization we consider $\lambda \in \{0, 0.5, 1\}$. Arrows with diamond endings show the feasible actions from each state. Rectangles identify the immediate rewards from each action. Directed arrows represent the stochastic outcomes of each action. Thick arrows with diamond endings indicate the educated policy for this contractor (i.e., the recommended career path).

5.1. The MDP Effect on Market Revenue and Worker Wages

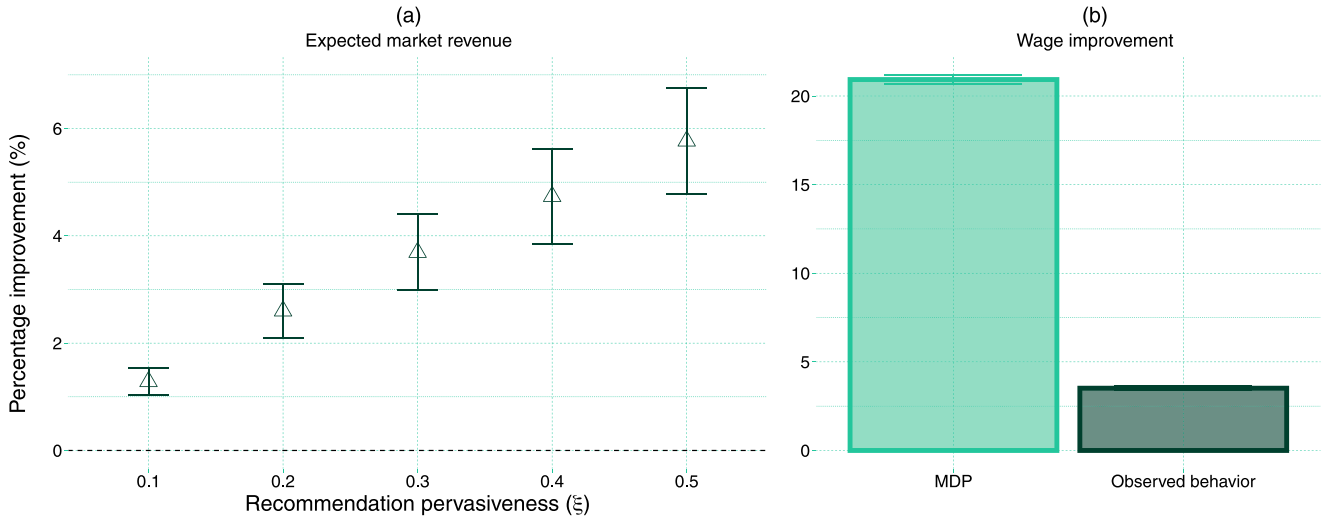
What would have happened if, at the end of year 2 of the focal data set, the marketplace decided to implement the MDP framework? To evaluate this scenario, we use the first two years to learn the parameters of the MDP framework, as well as the predictive models (XGBoost) for the demand, supply, and hourly wage (Equations (7)–(9)). We then split the final two years into two-month periods.⁴ At each period, contractors get a recommendation to work, learn, or work and learn a new set of relevant skills. Based on a “recommendation pervasiveness” parameter, $\xi \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$, contractors stochastically choose to follow the recommendation of the MDP framework. If contractors choose to learn a new skill, they stochastically transition to a new state according to their personalized transition probabilities (Equation (2)).

At the end of each period, we compute the changes in supply as a result of our recommendation framework. Based on this updated supply, we estimate the supply, demand, and hourly wages of the next period, according to Equations (7)–(9). We repeat this process for the $N = 12$ two-month periods of the last two years of the data set.

Our goal is to evaluate how the framework affects (1) the platform as a whole and (2) workers who follow its recommendations. To measure the effect on the platform, we estimate the mean expected hourly revenue of each period:

Expected market revenue:

$$I = \frac{1}{|\mathcal{S}| \cdot N} \sum_{\mathbf{S} \in \mathcal{S}} \times \sum_{n=1}^N [\hat{\text{Dem}}_n(\mathbf{S}) \cdot \hat{W}_n(\mathbf{S})]. \quad (11)$$

Figure 11. (Color online) Simulation Results on Real Data

Notes. Panel (a) shows the improvement of the MDP over the observed behavior in terms of expected market revenue. Panel (b) shows the wage improvement of the contractors who follow the MDP framework's recommendations compared with the observed contractor behavior during the same period. Error bars show 95% confidence intervals.

To measure the effect on workers who follow the recommendations, we estimate the per-period average wage improvement over the workers' prior wages:

$$\text{Wage improvement (\%)} = \frac{1}{N-1} \sum_{n=2}^N \frac{1}{|\mathbb{C}_n|} \sum_{i \in \mathbb{C}_n} \frac{\hat{W}_{in} - \hat{W}_{in-1}}{\hat{W}_{in-1}} * 100, \quad (12)$$

where \mathbb{C}_n is the set of contractors that successfully transition to a new state during period $n > 1$, and \hat{W}_{in} is contractor's i predicted wage at period n .

Figure 11(a) shows the results. The y -axis shows the average percentage improvement of implementing the MDP framework compared with the observed behavior of contractors during the last two years of the focal data set. For every value of the pervasiveness parameter ξ , the effect of the proposed framework would have been positive and statistically significant: the expected average increase in market revenue ranges between 1.7% and 6%. For a typical digital workplace with ~\$1B in annual transaction volume, a 6% revenue increase materializes to ~\$60 million.

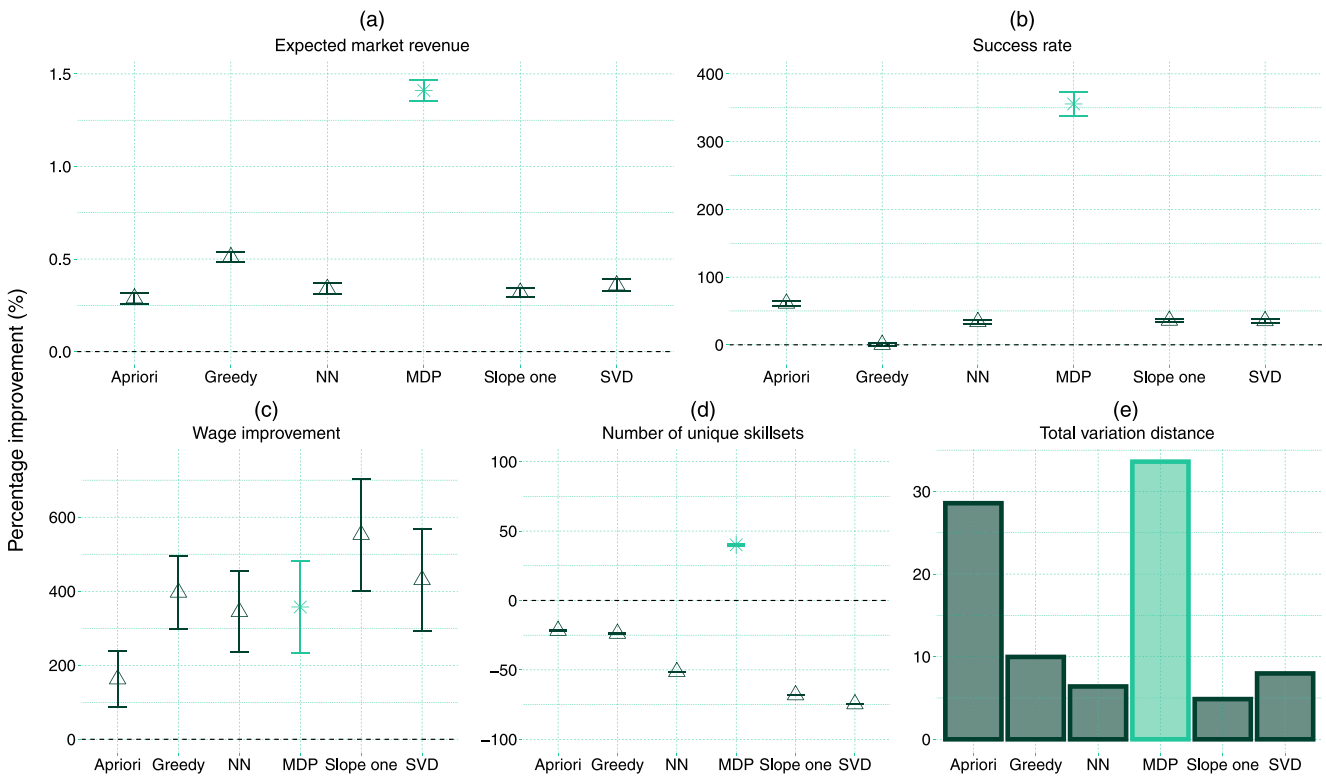
Figure 11(b) shows the wage improvement for contractors who transition to a new state (Equation (12)). The left column shows the results for contractors who follow the MDP recommendations, whereas the right column shows wage improvements for the observed contractor behavior during the same periods. On average, MDP-driven successful skill acquisitions yield a 22% wage increase. On the contrary, contractors currently make skill acquisitions that yield, on average, a wage improvement of only 3.5%.

5.2. Comparison of the MDP with Alternative Recommender Systems

The previous evaluation on observed behavior shows that implementing the proposed framework can be beneficial for both the market and the contractors who follow its recommendations. However, it does not compare how the MDP framework performs against alternative popular recommender systems. In this section, we benchmark the performance of the proposed MDP against six alternative recommender systems. Specifically, we model and evaluate the following approaches:

- Baseline:
 - *Current behavior*: This approach models the observed behavior of the contractors by recommending skills according to the observed distribution of skill acquisition (i.e., similar skills to similar contractors).
- Alternative recommender frameworks:
 - *Association rule mining (Apriori)*: Association rule mining (also known as basket analysis) discovers relations between "items" in large datasets. If we assume that a skillset is an "itemset" and a skill an "item," then the Apriori algorithm (Agrawal et al. 1994) can mine rules that recommend relevant skills given the current contractor skillset.⁵
 - *Collaborative filtering*: Section 2.3 argues that current recommender systems might not be suitable for our context, as they would rely on prior, often poor contractor choices to recommend career paths. Besides, because of context-specific peculiarities, these systems require various adaptations in order to recommend skills. Here, we consider three popular collaborative filtering frameworks: k -nearest

Figure 12. (Color online) MDP Performance Improvement over Alternative Models



neighbors (k NN; Kantor et al. 2011), singular value decomposition (SVD; Kantor et al. 2011), and slope one (Lemire and Maclachlan 2005). To implement these systems, we assume that each contractor is a “user,” each skill is an “item,” and contractors’ compensations are the “ratings.” Based on this matrix, we can make skill recommendations such as “contractors who have similar earnings and skillsets should also learn similar skills.”

◦ *Greedy*: This approach makes recommendations by suggesting feasible skills that give the highest immediate rewards (i.e., it does not optimize a long-term career path).

To evaluate these algorithms, we perform simulations on a set of 30,000 randomly generated contractors. These contractors are assigned learning abilities (d), countries (c), and states S according to the observed distributions of learning abilities, countries, and states in the focal data set. As a result, this set of contractors is distributionally identical to the observed set of contractors. Similar to before, we run simulations for two years ahead, and we use recommendation pervasiveness thresholds $\xi \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. To generate a fair, “all else being equal” comparison, each

algorithm simulates a parallel universe, where, in each universe, the same contractors follow different career paths according to the recommendations they receive.

We benchmark all approaches against the baseline (current behavior) recommender framework. Figure 12 shows the results. The y-axis shows how much better each approach performs compared with the baseline. Figure 12(a) shows that all frameworks yield higher market revenue than the baseline model. However, the proposed MDP framework yields significantly ($p < 0.001$) higher improvement over the baseline compared with any one of the alternative five recommender frameworks.

Figure 12(b) shows the success rates of the recommendations of each approach in comparison with the success rates of the baseline. The success rate captures the percentage of contractors who successfully learn a new skill given that they have been exposed to a skill recommendation. Recall that, for all approaches, contractors transition to a new state stochastically through Equation (2) (all else being equal environment). The MDP approach yields the highest success rate, which is four times higher than

the baseline, and significantly ($p < 0.001$) higher than any of the alternative models. This superiority of the MDP occurs because of the explicit incorporation of the contractors' characteristics and the learning affinities of the new skillsets in estimating the transition probabilities (Section 4.2). As a result, the MDP makes targeted, more personalized, and more relevant recommendations compared with alternative approaches.

Figure 12(c) compares all algorithms in terms of wage improvements for contractors who follow the algorithms' recommendations (Equation (12)). First, the figure shows that all models outperform the baseline ($p < 0.001$). This is because the baseline does not include any wage information in the proposed recommendations. Apriori, which also does not include wage information, ends up mining more relevant recommendations that yield higher-than-the-baseline wage improvements. However, all algorithms that include wage information in their recommendations perform significantly better than Apriori (at least $p < 0.1$) and the baseline ($p < 0.001$).

One could argue that the proposed framework (or any of the alternative frameworks) might end up recommending the same set of skills and saturate the market. Recall that, during the recommendation and the performance evaluation phases of the simulation, we use predictions of supply and demand to estimate the wage of the next period (Equations (7)–(9)). These predictions include changes in supply through the proposed recommendations. Based on the results presented in Sections 5.1 and 5.2, there is no evidence that the proposed recommendations saturate the market in a way that hurts contractor wages or the market's revenue.

To compare which of the seven models generates more diverse recommendations, Figure 12(d) illustrates the number of unique skillsets that each algorithm recommends. The proposed MDP framework recommends 47% more diverse skillsets than the baseline current behavior model. On the contrary, all five alternative models provide up to 75% less diverse recommendations than the baseline.

The number of unique recommended skillsets paints an incomplete picture of the diversity of recommendations, as it ignores the frequency by which each skillset is being recommended. Conceptually, a truly diverse recommender suggests all skillsets with similar frequencies. One way to measure this is to estimate how close the skillset frequency distributions are with the uniform distribution. The total variation distance (Huber 2011, Kohl 2019) between any two distributions captures how close these distributions are. Hence, we can use this to measure the closeness of the skillset frequency distributions with the uniform distribution.

Figure 12(e) shows that the proposed approach yields a skillset frequency distribution that is 34% closer to the uniform distribution compared with the baseline and much closer to the uniform distribution compared with the five alternative approaches.

This superiority of the MDP in terms of diversified recommendations is a direct result of its high level of stochasticity. In particular, stochasticity comes from (1) the learning ability of a contractor (d), (2) the percentage that a contractor chooses to work (λ), (3) the contractor country (c), and (4) the contractor state (S). Because of these, it is unlikely for the framework to recommend similar career paths, even when contractors have similar profiles.

In conclusion, these simulations show that the proposed framework can benefit both the contractors and the marketplace. In particular, the marketplace could experience an increase in revenue between 1.75% and 6%. Similarly, contractors who follow the system's recommendations successfully could receive, on average, a 22% increase in their wages. Comparison with six alternative approaches across five different evaluation metrics shows an overall superiority of the MDP approach, which is a result of the fact that the framework (1) does not rely on previous, often poor contractor choices to recommend future behavior, (2) takes into account current (and future) trends in demand, supply, and wages, and (3) offers diverse, personalized recommendations that match individual contractor characteristics. As a result, compared with alternative systems and the current contractor behavior, MDP recommendations yield higher expected market revenue, wage improvements, and success rates of skill acquisitions. At the same time, the stochasticity of the framework and the dynamic estimation of future wages yield more diverse skillset recommendations that virtually guarantee that the market will not saturate with contractors that own the same sets of skills.

6. Discussion

This work presents a personalized and adaptive career development robo-advisor. The framework combines concepts of reinforcement learning, deep learning, and Bayesian inference to provide recommendations of how contractors *should* behave in terms of acquiring new skills. Evaluation through simulations on a transactional data set from an online labor marketplace shows that an adaptation of the proposed framework could lead to (1) a 1.75%–6% increase of the market's revenue, (2) a 22% increase of compensation wages for contractors who follow the recommended career paths, and (3) an increased skillset diversity in the marketplace.

6.1. Contributions to Research and Managerial Implications

This study contributes to the design research of digital workplaces and has a series of managerial implications. In particular, the proposed framework resolves a series of shortcomings of current “vanilla” machine learning (and other) approaches, including (1) reliance on prior (often poor) contractor choices, (2) assumptions about nondynamic observed behaviors, and (3) lack of diversification. This is only feasible because the framework does not model previous behavior to recommend future behavior. Instead, it uses information from the market to identify feasible and likely profitable future career paths. Because it is built particularly for skill acquisitions, it further allows for the sequential evolution of contractors through learning, working, and learning while working. Finally, it leverages observed information to estimate a contractor’s learning ability and likelihood of a successful skill acquisition, which in turn results in stochastic recommendations that are significantly different between seemingly similar contractors. The result is a diversified, adaptive career development robo-advisor that models how contractors *should* behave in terms of skill acquisition.

The proposed framework can contribute to managerial decision-making in online and other workplaces. Because it operates on an up-to-date evaluation of the current state of the marketplace, it dynamically adapts to situations in which individual skills lose or gain value. These online adaptations increase the contractors’ awareness of the current (and future) state of the marketplace. Through better awareness, contractors can make better-informed decisions and increase their market value. Consequently, increased market value can lead to an increase in the employers’ satisfaction through a more efficient talent search process (Brynjolfsson and Smith 2000). This would, in turn, increase the overall transactional efficiency of the marketplace.

Even though our results are constrained within a given online labor market and a given set of skills, the methodology is widely applicable. For example, TaskRabbit could potentially use the proposed approach to identify the value of each available handyman skill and recommend relevant and profitable career adaptations. Professional networks such as LinkedIn could adapt this analysis to their context by associating the expected rewards of each action with promotions or new jobs. Matching platforms such as Monster or Hired could extract resume information to implement the proposed approach and recommend relevant career paths to their talent bases.

Our framework also applies in online—and given data availability in offline—educational platforms. Massive online open courses offered by platforms

such as Coursera, EDX, and Nanodegrees are growing fast (Adamopoulos 2013b). By gauging the demand and expected compensation of a given skill (through platforms such as Burning Glass), the proposed robo-advisor can recommend relevant clusters of online courses that can lead to successful career trajectories. Overall, and given that, by some estimates, 120 million workers worldwide will need to reeducate in the next three years (Institute of Business Value 2019), the proposed framework can find applications in diverse contexts and help workers make educated decisions regarding their career trajectories.

6.2. Limitations

One limitation of our approach is that it does not explicitly model the equilibrium effects of our reskilling suggestions. In particular, the estimation of demand and supply relies on predictive modeling, according to Equations (7) and (8). Through these equations and by accounting for workers’ reskilling decisions, our framework learns during training to respond to any supply effects of its recommendations (lines 11–16 of Algorithm 1). For instance, if several workers follow the framework’s recommendations at time t , the framework will observe changes in supply at time $t + 1$, and update its supply, demand, and wage predictions according to Equations (7), (8), and (9). Because this process happens during training (Algorithm 1), our approach learns a Q -function that anticipates supply shifts. Section 5 empirically shows that this approach significantly benefits the market and its workers. Regardless, future research can extend our method to include equilibrium effects by explicitly modeling the demand and supply of skills.

The prediction of future demand, supply, and wages creates a second limitation, as prediction errors propagate in the estimation of the Q -function (Algorithm 1, line 16). This is a necessary compromise for any framework that wants to account for future trends in supply and demand. In practice, platforms that implement our approach will provide up-to-date recommendations by continuously retraining the framework (e.g., once every month). Even further, with advances in machine learning, time-series prediction errors are relatively small (Appendix D, Figures 13 and 14); hence, such errors are unlikely to cancel out the benefits of the proposed approach (Section 5, Figure 11). In fact, Appendix F empirically shows that considering future supply and demand yields better results than learning a Q -function on a static environment (Figure 16).

Another limitation of our approach is that it requires a fixed set of available states \mathbb{S} in order to learn and recommend meaningful career paths. As we mentioned in Section 4.6, additional states may be possible but not yet observed in the marketplace

(i.e., no contractor has the skills to form these states). Besides, new skills constantly emerge in online labor markets. Set \mathbb{S} will miss such unobserved (or new) states, and the framework will not be able to recommend them. In practice, we can mitigate this limitation by adjusting the set of feasible states every time we update (retrain) the framework. For instance, if the framework trains once every month, it can update its set of feasible states \mathbb{S} monthly. In general, the framework can include newly observed states in its career path recommendations with a maximum delay equal to the time between consecutive trainings.

Finally, our approach implicitly ignores the contamination of the recommendations from potential communication between contractors. Such communication could lead to system mistrust, as contractors might perceive personalized recommendations as random. Even though this is possible, the online and global nature of the focal platform suggests that such effects will likely be limited. Besides, we believe that an implementation of the proposed system should be transparent and explain how the framework works (Datta et al. 2016, Rader et al. 2018). Algorithmic transparency regarding potential variations in recommendations will allow contractors to understand the framework's behavior and likely trust its recommendations.

6.3. Future Directions

An extension of the current framework could incorporate human input and become an augmented intelligence system (Jain et al. 2018). Workers could interact with the system by revealing how much time they intend to spend in learning or by providing accurate information about how much time they have spent in the past to acquire their current skillsets. Such manual inputs to the algorithm could remove noise from the framework's estimates and improve its performance.

Depending on the context, the components of the proposed framework can be refined and extended. The attached appendices provide various such extensions. For instance, Appendix A presents alternative approaches for decomposing skillsets to numeric vectors. Appendix B provides more advanced approaches for estimating contractors' learning abilities. Appendix E provides alternative approaches for action-space exploration, whereas Appendix F shows more advanced estimation processes that yield better results.

The proposed framework could further extend to consider the expertise of each contractor on the available skillsets. Appendix G presents a basic approach of how expertise estimates can structure the action-state space of Algorithm 1. Such an extension could capture transitions to subsets of current skillsets, as (1) contractors might lose the ability to perform

tasks on specific skills if they do not use them, or (2) some skills might lose value and become obsolete.

To conclude, this study identifies shortcomings of current recommender systems when applied to career development paths and proposes an appropriate, context-specific dynamic framework that resolves them. If deployed, the framework has the potential to generate positive effects for the marketplace and its contractors, as well as positive spillover effects on the marketplace employers.

Acknowledgments

The authors thank Sam Ransbotham, Panagiotis Adamopoulos, and Vilma Todri for their comments and suggestions on improving the paper.

Endnotes

¹ Results remain robust if we assume that the presence of a certification suggests a successful skill acquisition. This probably occurs because contractors tend to list certifications for which they have scored in the upper 50th percentile (Ipeirotis 2013).

² We use a threshold of 3.5/5 to separate successful from unsuccessful task outcomes. Results remain robust if we assume that any task completion signals a successful skill acquisition.

³ Equation (10) is an off-policy temporal difference control procedure (Sutton and Barto 1998, p. 107).

⁴ We implicitly assume that two months is enough time for a contractor to learn a new, relevant skill.

⁵ The choice of association rule mining is not ad hoc. During our collaboration with the focal market, association rule mining was the most discussed approach for estimating associations between different skillsets and implementing a skill recommendation framework.

References

- Adamopoulos P (2013a) Beyond rating prediction accuracy: On new perspectives in recommender systems. *Proc. 7th ACM Conf. Recommender Systems* (Association for Computing Machinery, New York), 459–462.
- Adamopoulos P (2013b) What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. *Internat. Conf. Inform. Systems* (AIS, Atlanta).
- Adamopoulos P (2014) On discovering non-obvious recommendations: Using unexpectedness and neighborhood selection methods in collaborative filtering systems. *Proc. 7th ACM Internat. Conf. Web Search Data Mining* (Association for Computing Machinery, New York), 655–660.
- Adamopoulos P, Tuzhilin A (2014) On over-specialization and concentration bias of recommendations: Probabilistic neighborhood selection in collaborative filtering systems. *Proc. 8th ACM Conf. Systems* (Association for Computing Machinery, New York), 153–160.
- Adamopoulos P, Tuzhilin A (2015a) The business value of recommendations: A privacy-preserving econometric analysis. *Internat. Conf. Inform. Systems* (AIS, Atlanta).
- Adamopoulos P, Tuzhilin A (2015b) On unexpectedness in recommender systems: Or how to better expect the unexpected. *ACM Trans. Intelligent Systems Tech.* 5(4):1–32.
- Adamopoulos P, Tuzhilin A, Mountanos P (2015) Measuring the concentration reinforcement bias of recommender systems. *9th ACM Conf. Recommender Systems* (Association for Computing Machinery, New York).

- Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Trans. Knowledge Data Engrg.* 17(6):734–749.
- Agarwal R, Ohyama A (2013) Industry or academia, basic or applied? Career choices and earnings trajectories of scientists. *Management Sci.* 59(4):950–970.
- Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. *Proc. 20th Internat. Conf. Very Large Data Bases* (Morgan Kaufmann Publishers, Burlington, MA), 487–499.
- Agrawal A, Horton J, Lacetera N, Lyons E (2015) Digitization and the contract labor market. *Economic Analysis of the Digital Economy*, 219.
- Akerlof GA (1970) The market for “lemons”: Quality uncertainty and the market mechanism. *Quart. J. Econom.* 84(3):488–500.
- Allon G, Bassamboo A, Çil EB (2012) Large-scale service marketplaces: The role of the moderating firm. *Management Sci.* 58(10):1854–1872.
- Arora A, Forman C (2007) Proximity and information technology outsourcing: How local are it services markets? *J. Management Inform. Systems* 24(2):73–102.
- Autor DH (2001) Wiring the labor market. *J. Econom. Perspect.* 15(1):25–40.
- Autor DH, Katz LF, Krueger AB (1998) Computing inequality: Have computers changed the labor market? *Quart. J. Econ.* 113(4): 1169–1213.
- Balabanović M, Shoham Y (1997) Fab: Content-based, collaborative recommendation. *Commun. ACM* 40(3):66–72.
- Banker RD, Hwang I (2008) Importance of measures of past performance: Empirical evidence on quality of e-service providers. *Contemporary Accounting Res.* 25(2):307–337.
- Becker GS (1962) Investment in human capital: A theoretical analysis. *J. Political Econom.* 70(5):9–49.
- Becker GS (1994) Human capital revisited. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, 3rd ed. (University of Chicago Press, Chicago), 15–28.
- Bennett J, Lanning S (2007) The Netflix prize. *Proc. Internat. Conf. Knowledge Discovery Data Mining Cup Workshop* (Association for Computing Machinery, New York), 35.
- Berinsky AJ, Huber GA, Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon.com’s Mechanical Turk. *Political Anal.* 20(3):351–368.
- Bian L, Holtzman H (2011) Online friend recommendation through personality matching and collaborative filtering. *Internat. Conf. Mobile Ubiquitous Comput. Systems Services Tech.* (IARIA, Wilmington, DE), 230–235.
- Billsus D, Pazzani MJ (1998) Learning collaborative information filters. Shavlik JW, ed. *Proc. 15th Internat. Conf. Machine Learn.* (Morgan Kaufmann, Burlington, MA), 46–54.
- Blundell R, Dearden L, Meghir C, Sianesi B (1999) Human capital investment: The returns from education and training to the individual, the firm and the economy. *Fiscal Stud.* 20(1):1–23.
- Bosma N, Van Praag M, Thurik R, De Wit G (2004) The value of human and social capital investments for the business performance of startups. *Small Bus. Econom.* 23(3):227–236.
- Breese SJ, Heckerman D, Kadie C (1998) Empirical analysis of predictive algorithms for collaborative filtering. *Proc. 14th Conf. Uncertainty Artificial Intelligence* (Morgan Kaufmann Publishers, Burlington, MA), 43–52.
- Brier E, Pearson R (2018) Upwork’s SVP of marketing explains what it takes to perfect an offering that relies on people. Accessed April 28, 2019, <https://www.prnewswire.com/news-releases/snagajob-appoints-former-upwork-ceo-to-board-of-directors-300417689.html>.
- Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* 46(4):563–585.
- Brynjolfsson E, Hu Y, Simester D (2011) Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Sci.* 57(8):1373–1386.
- Chen DL, Horton JJ (2016) Are online labor markets spot markets for tasks? A field experiment on the behavioral response to wage cuts. *Inform. Systems Res.* 27(2):403–423.
- Claypool M, Gokhale A, Miranda T, Murnikov P, Netes D, Sartin M (1999) Combining content-based and collaborative filters in an online newspaper. *Proc. ACM SIGIR Workshop Recommender Systems* (Association for Computing Machinery, New York).
- Cooper AC, Woo CY, Dunkelberg WC (1988) Entrepreneurs’ perceived chances for success. *J. Bus. Venturing* 3(2):97–108.
- Datta A, Sen S, Zick Y (2016) Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. *Sympos. Security Privacy* (IEEE, Piscataway, NJ), 598–617.
- Delgado J, Ishii N (1999) Memory-based weighted majority prediction. *Special Interest Group Inform. Retrieval Workshop Recommender Systems* (Association for Computing Machinery, New York).
- Dey D, Fan M, Zhang C (2010) Design and analysis of contracts for software outsourcing. *Inform. Systems Res.* 21(1):93–114.
- Fischhoff B, Slovic P, Lichtenstein S (1977) Knowing with certainty: The appropriateness of extreme confidence. *J. Experiment. Psych. Human Perceptions Performance* 3(4):552–564.
- Fleder D, Hosanagar K (2009) Blockbuster culture’s next rise or fall: The impact of recommender systems on sales diversity. *Management Sci.* 55(5):697–712.
- Freelancers Union, Upwork (2017) Freelancing in America. Accessed December 2, 2019, https://s3-us-west-1.amazonaws.com/adquiro-content-prod/documents/Infographic_UP-URL_2040x1180.pdf.
- Gati I, Levin N (2014) Counseling for career decision-making difficulties: Measures and methods. *Career Development Quart.* 62(2): 98–113.
- Gefen D, Carmel E (2008) Is the world really flat? A look at offshoring at an online programming marketplace. *Management Inform. Systems Quart.* 32(2):367–384.
- Getoor L, Sahami M (1999) Using probabilistic relational models for collaborative filtering. *Workshop Web Usage Anal. User Profiling* (Springer, Verlag, Berlin Heidelberg).
- Glomm G, Ravikumar B (1992) Public vs. private investment in human capital: endogenous growth and income inequality. *J. Political Econom.* 100(4):818–834.
- Gomez-Urbe CA, Hunt N (2016) The Netflix recommender system: Algorithms, business value, and innovation. *Trans. Management Inform. Systems* 6(4):1–19.
- Gopal A, Sivaramakrishnan K, Krishnan MS, Mukhopadhyay T (2003) Contracts in offshore software development: An empirical analysis. *Management Sci.* 49(12):1671–1683.
- Horton JJ (2010) Online labor markets. Saberi A, ed. *Internet and Network Economics*, Lecture Notes in Computer Science, vol. 6484 (Springer, Berlin, Heidelberg), 515–522.
- Horton JJ, Chilton LB (2010) The labor economics of paid crowdsourcing. *Conf. Electronic Commerce* (Association for Computing Machinery, New York), 209–218.
- Horton JJ, Rand DG, Zeckhauser RJ (2011) The online laboratory: Conducting experiments in a real labor market. *Experiment. Econom.* 14(3):399–425.
- Hu J (2016) Reinforcement learning explained. Accessed December 30, 2018, <https://www.oreilly.com/ideas/reinforcement-learning-explained>.
- Huber PJ (2011) *Robust Statistics* (Springer, Berlin, Heidelberg).
- Institute of Business Value (2019) The enterprise guide to closing the skills gap. Accessed December 2, 2019, <https://www.ibm.com/downloads/cas/EPYMNBJA>.
- Ipeirotis P (2013) Badges and the Lake Wobegon effect. Accessed December 30, 2019, <https://www.behind-the-enemy-lines.com/2013/10/badges-and-lake-wobegon-effect.html>.

- Ipeirotis PG, Provost F, Wang J (2010) Quality management on Amazon Mechanical Turk. *Workshop Human Comput.* (Association for Computing Machinery, New York), 64–67.
- Jain H, Padmanabhan B, Pavlou PA, Santanam RT (2018) Call for papers—special issue of information systems research—humans, algorithms, and augmented intelligence: The future of work, organizations, and society. *Inform. Systems Res.* 29(1):250–251.
- Johnson DDP, Fowler JH (2011) The evolution of overconfidence. *Nature* 477(7364):317–320.
- Johnson T (1970) Returns from investment in human capital. *Amer. Econom. Rev.* 60(4):546–560.
- Kantor BP, Rokach L, Ricci F, Shapira B (2011) *Recommender Systems Handbook* (Springer, New York).
- Kohl M (2019) Totalvardist: Generic function for the computation of the total variation distance of two distributions. Accessed December 2, 2019, <https://www.rdocumentation.org/packages/distrEx/versions/2.5/topics/TotalVarDist>.
- Kokkodis M (2018) Dynamic recommendations for sequential hiring decisions in online labor markets. *Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 453–461.
- Kokkodis M (2019) Reputation deflation through dynamic expertise assessment in online labor markets. *World Wide Web Conf.* (ACM), 896–905.
- Kokkodis M (2020) Skillset Diversification in Online Labor Markets: Reputation Losses and Opportunity Gains. Working paper, Boston College, Boston.
- Kokkodis M, Ipeirotis PG (2013) Have you done anything like that? Predicting performance using inter-category reputation. *Internat. Conf. Web Search Data Mining* (ACM, New York), 435–444.
- Kokkodis M, Ipeirotis PG (2014) The utility of skills in online labor markets. *Internat. Conf. Inform. Systems* (Association for Information Systems, Atlanta).
- Kokkodis M, Ipeirotis PG (2016) Reputation transferability in online labor markets. *Management Sci.* 62(6):1687–1706.
- Kokkodis M, Ipeirotis PG (2020) The good, the bad, and the unhirable: Recommending job applicants in online labor markets. Working paper Boston College, Boston.
- Kokkodis M, Papadimitriou P, Ipeirotis PG (2015) Hiring behavior models for online labor markets. *Proc. Eighth ACM Internat. Conf. Web Search Data Mining* (Association for Computing Machinery, New York), 223–232.
- Kuhn P, Skuterud M (2004) Internet job search and unemployment durations. *Amer. Econom. Rev.* 94(1):218–232.
- Lam XN, Vu T, Le TD, Duong AD (2008) Addressing cold-start problem in recommendation systems. *Proc. 2nd Internat. Conf. Ubiquitous Inform. Management Comm.* (ACM, New York), 208–211.
- Lang K (1995) Newsweeder: Learning to filter netnews. Friedl A, Russell S, eds. *Internat. Machine Learn. Conf.* (Morgan Kaufmann, Burlington, MA), 331–339.
- Lauren D (2017) Snagajob appoints former Upwork CEO to board of directors. Accessed April 28, 2019, <https://www.prnewswire.com/news-releases/snagajob-appoints-former-upwork-ceo-to-board-of-directors-300417689.html>.
- Lazear EP (2009) Firm-specific human capital: A skill-weights approach. *J. Political Econom.* 117(5):914–940.
- Lee W-C, Cheng B-W (2008) Effects of using a portable navigation system and paper map in real driving. *Accident Anal. Prevention* 40(1):303–308.
- Lemire D, Maclachlan A (2005) Slope one predictors for online rating-based collaborative filtering. *Internat. Conf. Data Mining* (SIAM, Philadelphia), 471–475.
- Lepak DP, Snell SA (1999) The human resource architecture: Toward a theory of human capital allocation and development. *Acad. Management Rev.* 24(1):31–48.
- Liang C, Hong Y, Gu B (2016) Effects of IT-enabled monitoring systems in online labor markets. *Internat. Conf. Inform. Systems* (AIS, Atlanta).
- Lin M, Liu Y, Viswanathan S (2016) Effectiveness of reputation in contracting for customized production: Evidence from online labor markets. *Management Sci.* 64(1):345–359.
- Marshall A (2009) *Principles of Economics: Unabridged*, 8th ed. (Cosimo Inc., New York).
- Mason W, Watts DJ (2010) Financial incentives and the performance of crowds. *SIGKDD Explorations Newsletter* 11(2):100–108.
- Melville P, Mooney RJ, Nagarajan R (2002) Content-boosted collaborative filtering for improved recommendations. *18th National Conf. Artificial Intelligence* (ACM, New York), 187–192.
- Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. Preprint, submitted January 16, <https://arxiv.org/abs/1301.3781>.
- Mooney RJ, Roy L (2000) Content-based book recommending using learning for text categorization. *Conf. Digital Libraries* (ACM, New York), 195–204.
- Moreno A, Terwiesch C (2014) Doing business with strangers: Reputation in online service marketplaces. *Inform. Systems Res.* 25(4):865–886.
- Mustafa N, Ibrahim AO, Ahmed A, Abdullah A (2017) Collaborative filtering: Techniques and applications. *Internat. Conf. Comm. Control Comput. Electronics Engrg.* (IEEE, Piscataway, NJ), 1–6.
- Myers DG (2011) *Psychology in Everyday Life* (Macmillan, New York).
- Odean T (1998) Volume, volatility, price, and profit when all traders are above average. *J. Finance* 53(6):1887–1934.
- Oliver B (2015) Redefining graduate employability and work-integrated learning: Proposals for effective higher education in disrupted economies. *J. Teaching Learn. Graduate Employability* 6(1):56–65.
- Pathak B, Garfinkel R, Gopal RD, Venkatesan R, Yin F (2010) Empirical analysis of the impact of recommender systems on sales. *J. Management Inform. Systems* 27(2):159–188.
- Pazzani M, Billsus D (1997) Learning and revising user profiles: The identification of interesting web sites. *Machine Learn.* 27(3):313–331.
- Pazzani MJ (1999) A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Rev.* 13(5-6):393–408.
- Phelan O, McCarthy K, Smyth B (2009) Using Twitter to recommend real-time topical news. *Proc. Third ACM Conf. Recommender Systems* (ACM, New York), 385–388.
- Rader E, Cotter K, Cho J (2018) Explanations as mechanisms for supporting algorithmic transparency. *Proc. Conf. Human Factors Comput. Systems* (ACM, New York), 1–13.
- Rand DG (2012) The promise of mechanical turk: How online labor markets can help theorists run behavioral experiments. *J. Theoret. Biol.* 299:172–179.
- Ricci F, Rokach L, Shapira B (2011) Introduction to recommender systems handbook. *Recommender Systems Handbook* (Springer, New York), 1–35.
- Rosen S (1983) Specialization and human capital. *J. Labor Econom.* 1(1):43–49.
- Sahoo N, Singh PV, Mukhopadhyay T (2012) A hidden Markov model for collaborative filtering. *Management Inform. Systems Quart.* 36(4):1329–1356.
- Schafer JB, Konstan J, Riedl J (1999) Recommender systems in e-commerce. *Proc. 1st ACM Conf. Electronic Commerce* (ACM, New York), 158–166.
- Schultz C (2018) Upwork press release. Accessed April 28, 2019, <https://www.upwork.com/press/2018/05/01/q1-2018-skills-index/>.
- Schultz TW (1961) Investment in human capital. *Amer. Econom. Rev.* 51(1):1–17.
- Sheng VS, Provost F, Ipeirotis PG (2008) Get another label? Improving data quality and data mining using multiple, noisy labelers.

- Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 614–622.
- Si L, Jin R (2003) Flexible mixture model for collaborative filtering. *Proc. 20th Internat. Conf. Machine Learn.* (AAAI, Palo Alto, CA), 704–711.
- Smith GE, Venkatraman MP, Dholakia RR (1999) Diagnosing the search cost effect: Waiting time and the moderating impact of prior category knowledge. *J. Econom. Psych.* 20(3):285–314.
- Stevenson B (2009) The internet and job search. Autor DH, ed. *Studies of Labor Market Intermediation* (University of Chicago Press, Chicago), 67–86.
- Sundararajan A (2016) *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism* (MIT Press, Cambridge, MA).
- Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge).
- Toma CL, Hancock JT, Ellison NB (2008) Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Perspect. Soc. Psych. Bull.* 34(8):1023–1036.
- Tso-Sutter HLK, Marinho LB, Schmidt-Thieme L (2008) Tag-aware recommender systems by fusion of collaborative filtering algorithms. *Proc. 2008 ACM Sympos. Appl. Comput.* (Association for Computing Machinery, New York), 1995–1999.
- Upwork (2014) Online work report. Accessed April 28, 2019, <https://web.archive.org/web/20180228011632/http://elance-odesk.com:80/online-work-report-global>.
- Upwork Enterprise (2017) Quarterly trends report. Accessed April 28, 2019, <https://www.upwork.com/s/enterprise-trends/>.
- Watkins CJCH, Dayan P (1992) Q-learning. *Machine Learn.* 8(3–4): 279–292.
- Wiltbank R, Read S, Dew N, Sarasvathy SD (2009) Prediction and control under uncertainty: Outcomes in angel investing. *J. Bus. Venturing* 24(2):116–133.
- Yoganarasimhan H (2013) The value of reputation in an online freelance marketplace. *Marketing Sci.* 32(6):860–891.