Spring 5-6-2024

Towards Equitable Urban Policies

Henry T. Woodyard VI
Georgia State University

ABSTRACT

TOWARDS EQUITABLE URBAN POLICIES

By

HENRY THOMAS WOODYARD VI

MAY, 2024

Committee Chair: Dr. Carlianne Patrick

Major Department: Economics

In recent years, increasing attention has been given to the role of space in not only generating inequality but also alleviating it through place-based policies. This research focuses on mechanisms for inequality generation through local labor markets and potential routes for mitigation. In “Skills, Matching, and Skill Specificity Across Space”, we test whether urban agglomeration is skill-biased by using text data on skills from a near universe of job postings and resumes. Creating a new measure of skill specificity by modeling the network of relationships between skill, we find evidence that an increase in urban population increases match quality on average and the premium is greater for specific skills. Premiums appear to be driven by both labor market thickness and sorting between cities.

Early childhood education is often regarded as an ideal economic development investment. Numerous studies on high-quality, model programs in the 1960s and 1970s demonstrated a strong link between participation in pre-K programs and both short-term student achievement and positive later-life outcomes. However, evidence on state-funded, ‘universal’, pre-K programs is inconclusive. In “Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia”, we use enrollment lotteries for over-subscribed school-based sites in Georgia’s Pre-K Program to analyze the impact of participation on

elementary school outcomes. Lottery winners enter kindergarten more prepared in both math and reading, but gains fade by the end of kindergarten. Further, some negative achievement effects emerge by grade 4. Our evidence suggests greater benefits and lesser attenuation of gains for economically disadvantaged students.

Monopsony power in the labor market drives a substantial portion of between-city wage inequality by allowing firms in smaller areas to set wages below the competitive wage. In "Monopsony in the Market for Remote Work", I use a double machine-learning estimator and data on job postings in the United States between 2012 and 2022 to estimate the elasticity of labor supply for fully-remote jobs. The small estimated elasticity indicates the presence of monopsony power in the market for fully-remote work. Furthermore, differences in fully-remote work's monopsony power across city size persists despite the work's geographically divorced nature.

TOWARDS EQUITABLE URBAN POLICIES

BY

HENRY THOMAS WOODYARD VI

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2024

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Dr. Carlianne Patrick

Committee: Dr. Dan Kreisman
Dr. Cathy Liu
Dr. Tim Sass

Electronic Version Approved:

Thomas J. Vicino, Dean
Andrew Young School of Policy Studies
Georgia State University
May, 2024

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Carlianne Patrick, for her excellent guidance in both constructing a dissertation and professionally developing as an economist. She gave me the freedom to try new ideas and the result is the research contained herein. I would also like to thank my committee members – Dr. Dan Kreisman, Dr. Tim Sass, and Dr. Cathy Liu – for their help in my writing of this dissertation and their excellent suggestions for the improvement of this research. Further, I must thank my family and friends for their support during the ups and downs of this process. To all of you: I could not have done this without your help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The interaction between space and inequality cannot be ignored. Space is a key driver of inequality, with one-third of the increase in wage inequality across the United States since the 1980s due to city size (Baum-Snow & Pavan, 2013). These increases in inequality have garnered widespread interest within both the academic and policy spheres. Place-based policies, which aim to develop specific local economies, have recently gained support as a way to redistribute wealth across space (Gaubert, Kline, & Yagan, 2020) and lift up economically depressed places (Bartik, 2015). Despite this resurgence, academic consensus on the efficiency of place-based policies has not yet been reached, especially compared to investments like education which have demonstrated returns (Jackson et al., 2016).

My research agenda focuses on inequality generation through local labor markets and potential routes for mitigation. The first essay seeks to determine the mechanism that causes the urban agglomeration premium to be skill-biased, a key question for understanding the role of cities in generating inequality. The second essay examines whether Georgia's universal pre-kindergarten program, an investment often raised as an alternative to place-based ones, generates measurable and persistent outcomes for students. Finally, the third essay tests for the existence of monopsony power in remote labor markets, helping us better understand wage setting in the work-from-home environment.

Skill-biased agglomeration economies explain about 80% of the urban contribution to wage inequality (Baum-Snow, Freedman, & Pavan, 2018). In "Skills, Matching, and Skill Specificity Across Space", I test whether urban size increases the match quality between skills supplied and demanded, and whether this relationship depends on how specific a skill is. I use skills extracted from a near-universe of job postings and worker profiles across the United States over the past ten years. I then construct a network from the skills in documents and use network centrality measures to define specificity. In short, I find that skills supplied by workers and demanded by firms are better matched in larger markets, and this effect is stronger for high specificity skills. These results help in rationalizing skill-biased agglomeration in the framework of the matching micro-foundation (Duranton & Puga, 2004).

Local policymakers often seek to make investments into the community through educational spending. Indeed, funding K-12 education tends to take the largest portion of each state's budget, and investments into pre-kindergarten education in particular have been steadily increasing (Francis & Randall, 2017). Despite this, evidence on the benefits of state-funded "universal" pre-K programs has been mixed (Huizen & Plantenga, 2018). In "Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia", I compare the elementary school outcomes of students who won a lottery for attending an oversubscribed school-based pre-K to students who lost. While several empirical limitations are present, I show evidence that boosts to academic outcomes from pre-K attendance fade quickly, mirroring recent findings for Tennessee's Voluntary Pre-K Program (Durkin, Farrin, & Wiesen, 2022). However, I do find larger, more persistent effects for low-income students.

Monopsony in the labor market occurs when that market has only a single employer of labor. More commonly, markets will have varying levels of monopsony *power*, a measure describing the ability of employers to set wages lower than would prevail in a competitive market (Boal & Ransom, 1997). Practically, this metric is retrieved by estimating the elasticity of labor supply in a market, with monopsonistic markets having inelastic labor supply. Monopsony power tends to decrease with labor market size, with small markets having nearly inelastic labor supply (Azar, Marinescu, & Steinbaum, 2019), and may explain about 20% of the wage disparity between large and small markets (Luccioletti, 2023). Recently, Dube, Jacobs, Naidu, and Suri (2020) found not-negligible monopsony power in the online labor market Amazon Mechanical Turk (MTurk), despite having many traditional characteristics of competitive markets. Meanwhile, the availability of remote work in the traditional labor market continues to rise. In "Monopsony in the Market for Remote Work", I reproduce Dube et. al (2020)'s methodology with job postings in the United States in order to measure monopsony power in the US remote labor market. I find evidence of monopsony power in remote markets and, despite the geographically-divorced nature of the work, find variation in this monopsony power across space. These results inform the extent to which remote work can be welfare enhancing, especially for workers in small, monopsonistic markets.

CHAPTER I. Skills, Matching, and Skill Specificity Across Space

(with Carlianne Patrick)

1.1. Introduction

Place-based policies must be rooted in communities' labor market conditions to achieve inclusive growth. Growing evidence of shifting demand and supply for skills across space (e.g., Moretti, 2012; Diamond, 2016; Weinstein & Patrick 2020), which has been further linked to the rise in inequality (Florida et al., 2012; Austin et al., 2018; Giannone, 2018; Baum-Snow et al., 2019; Autor 2019). For example, one-third of the increase in wage inequality across the United States since the 1980s can be attributed just to city size (Baum-Snow & Pavan, 2013). About 80% of this increase is due to the skill bias of agglomeration, wherein higher skilled workers see greater returns to productivity and wages from urban density (Baum-Snow et al., 2019). Most of this research relies on proxying skill using other measures like educational attainment or occupational skills. Yet, skills vary widely within educational attainment group (Autor et al., 2003; Wolff, 2003; Ingram & Neumann, 2006; Heckman et al., 2006; Poletaev & Robinson, 2008; Bacolod et al., 2009a & 2009b; Zlatko & Ajwad, 2014; Weinstein & Patrick 2020). Furthermore, research using occupational definitions and employment relies on equilibrium outcomes and does not account for spatial variation in skills within occupations.

This paper contributes to the literature on spatial variation in the supply and demand of skills, inequality within and across space, and the microfoundations of agglomeration externalities with several important improvements over existing work.

First, we identify skills directly from text in job postings and resumes rather than using proxies such as educational attainment. Recent work by Atalay et al. (2021) uses similar data to identify job tasks and documents the variation in job task requirements across locations (and within occupational groups across locations) from postings. Our work is complementary, but distinct in that we focus on skills rather than tasks in job postings and resumes. Skills are anything that defines someone's knowledge, experience, and abilities. They are worker-oriented and common to both workers and jobs. Tasks, on the other hand, are job-oriented activities. A skill may be applicable to many tasks and vice versa. Importantly, skills are transferable across tasks.

Second, we develop a measure of skill specificity using novel data processing methods in which we take advantage of the relationships between skills and generate "skill networks." In the network, the nodes are skills listed in resumes. Edges connect skill nodes if the two skills are observed in a resume together, and the edges are weighted by the number of times the two appear together. For each skill, we calculate the closeness centrality, which is the measure of average weighted distance to all other nodes in the graph. Using this, we call skills that are less central more specific.

Importantly, our novel method allows us to define specificity at the skill level rather than at the occupation, job, or firm-market level. Previous work uses notions of scarcity or specialization that compare the attribute (e.g., tasks, etc.) vectors between two jobs or occupations using methods to characterize the degree of (non-) overlap, e.g. cosine dissimilarity. The overlap measures are then aggregated to describe the degree of specialization at the job, occupation, or firm-market level. Instead, our measure uses information on the network of relationships within the entire set of job postings and resumes to define the specificity at the skill level. Using our skill specificity measure, we document the spatial variation in the supply and demand of skills along the specificity distribution. We find that demand for skill specificity increases with labor market size.

Our third innovation exploits the skill detail in resumes and job-postings to develop two measures of skill mismatch within locations. The first method simply compares the percentage difference in the number of times the skill appears in resumes in that city versus job postings in that city. Because this approach defines mismatch at the skill level, it allows us to segment the analysis over the distribution of skill specificity. The second approach to defining match quality in a location simulates a local labor market in which unemployed workers and hiring firms are matched based on the distance between skills listed on resumes and job postings. Our measures of match quality allow us to consider both supply and demand, assessing differences within markets that have heretofore been undocumented.

Finally, we regress skill mismatch (along the skill specificity distribution) on locational characteristics. To provide causal estimates and control for the endogeneity of some locational characteristics (e.g., match quality and city size, etc.), we follow previous work and use standard instruments for these characteristics (i.e., geological and lagged population for city size and density). Across all methods, models, and samples, we find consistent evidence that an increase in urban population increases match quality on average. Further, the premium is greater for skills that are more specific, giving evidence that the matching microfoundation has a role in the generation of skill-biased agglomeration economies. These effects stay relatively constant when the sample is restricted to only frequent skills, and effects are significantly higher when restricted to only job postings requiring a bachelor’s degree or higher.

1.2. Background

Space and skill are strongly intertwined. Since the 1980s, higher-educated workers have been increasingly choosing to live and work in urban areas, where they see greater productivity, wages, and amenities (Diamond, 2016). This sorting may be facilitated by relatively lower mobility costs and spatial frictions for white-collar workers (Schmutz et al., 2019). As this sorting happens, local productivity increases endogenously, further increasing wages (Diamond, 2016; Baum-Snow et al., 2018). Indeed, of the one-third of inequality increase since the 1980s attributable to city size (Baum-Snow & Pavan, 2013), 80% of it is explained by skill-biased productivity gains from urban agglomeration (Baum-Snow et al., 2018) wherein urban growth yields wage or productivity premiums from external economies of scale.

Urban agglomeration externalities are best characterized by Duranton and Puga (2004), who present three possible driving mechanisms, or micro-foundations: sharing, learning, and matching. “Sharing” describes the pooling of, for example, input markets or risk; “learning” describes information spillovers between proximal workers and firms; and “matching” describes decreased mismatch between the skills provided by workers and demanded by firms. The focus of this paper is the third. Duranton and Puga present a model of agglomeration in which workers with heterogeneous skills match with a firm subject to a given wage and a cost of skill mismatch. The mismatch cost reflects the necessity to train relevant skills to reduce mismatch, which is passed through to workers in the form of lower wages. In this model, as firms enter, the average worker is able to find a better match for their skills. This has an important implication for workers: as the size of the city increases, the average worker is able to find a better match, become more productive more quickly, and be paid a higher wage.

Empirically, the methods previous authors have used to test this hypothesis have varied. Andersson, Burgess, and Lane (2007), using linked individual-firm data from the LEHD, define worker and firm quality as the respective coefficients of the individual and firm fixed-effects in a wage regression. That paper presents evidence that matching between high quality workers and firms is stronger in denser areas. In addition, that study sees strong complementarities in production between high quality workers and firms. Dauth et al. (2022), using individual-level linked worker-firm data from Germany, takes the approach of Andersson, Burgess, and Lane (2007) in defining worker and firm quality. They confirm the findings of Andersson, Burgess, and Lane (2007), finding that better matching between high- and low-quality workers in larger cities generates an urban wage premium. Furthermore, their results indicate that the strength of this type of matching has increased over time.

Another common strategy attempts to use past information about individuals to infer the position they would best be suited for. Harmon (2013) uses past industry experience to measure the quality of the current match. He determines that individuals in larger labor markets are more likely to find higher quality matches, and that subsequently they see higher earnings and are more likely to be retained. Abel and Deitz (2015) measure match quality along two dimensions: first, the level of education an occupation requires, and second, the specific college major with which an occupation is associated. Their results indicate that workers in thicker labor markets are more likely to find a job related to their college degree, and that well-matched workers see a significant wage premium.

It is possible that the benefit to matching from density works through improving search and job mobility. Indeed, more firms and workers trying to match improves the probability of matching (Petrongolo & Pissarides, 2001). Using data from the NLSY, Wheeler (2008) provides evidence that the probability of a worker changing industries increases with the size of the labor market for younger workers yet decreases for older workers. Di Addario (2011) show that the probability of finding a job conditional on searching increases with the size of the local labor market, with no commensurate changes to search intensity. City size does not appear to decrease the time to find a match, however (Harmon, 2013).

More recent studies that assess match quality have used scores from military aptitude tests to measure talents. Fredriksson, Hensvik, and Skans (2018) combine administrative data from Sweden with test data from the Swedish War Archives to observe individual occupational histories as well as scores in four subtests: reasoning, verbal comprehension, spatial ability, and technical understanding. Under the assumption that tenured workers in a given firm are well-matched, that paper compares new employees to incumbent workers along the dimensions of the test to determine mismatch. They find that mismatch affects inexperienced workers relatively more than experienced ones, generating downward pressure on wage growth and increased separations. Interestingly, they find evidence that about half of sorting happens within the same occupational classification, as workers sort according to their aptitudes. Guvenen et al. (2020) take a similar approach using individual-level data with ASVAB scores from the NLSY. Mismatch is found to have a persistent negative impact on wages even after moving to a new occupation, which the author argues is a result of reduced skill acquisition during the mismatch.

The theme which naturally emerges from this description of the literature is the variety of methods authors have used to measure what it means to be "high-skilled" and, likewise, what it means to be mismatched. Methods for measuring skill discussed so far have included educational attainment, implied ability (estimated via fixed effects), aptitude test, et cetera. Such a variety of methods indicate either the uncertainty in what it means to be "skilled", empirical challenges in measuring it, or both.

1.3. Skills Data

The skills data is obtained from job postings and worker profiles scraped from online job boards purchased from Emsi, Inc.[1] The data provides a rich overview of job vacancies and workers in the United States. It is large, containing more than 321 million job postings between 2010 and 2022 and almost 132 million profiles; for reference, according to the BLS, there were about 162 million people employed in the US in December of 2021. Emsi scrapes job postings and profiles from company websites, job boards, and aggregators, after which the raw text is parsed using natural language processing. Their algorithm parses out salient information for each category of document. For postings, these include the company name, job title, location, degree requirements, experience requirements, industry, and occupation, as well as indicators for remote work, part-time employment, and internships. The profiles data includes the current location, occupation, and industry of an individual, as well as their posted skills, educational history, and employment history. The educational history provides the name of the institution, the IPEDS code, degree level, and major.

In 2019, Emsi created a dictionary defining 32,000 skills and an algorithm to extract them from unstructured text, generating a list of skills included in each posting and profile. Over four billion skills in total are extracted from job postings alone. Table 1 presents basic information regarding the extracted skills. The average worker profile has roughly fourteen skills, while the average posting has about thirteen.

[1] Emsi, Inc. merged with Burning Glass Technologies in 2022 to become Lightcast.

Table 1. Basic Information on Skills in Profiles and Postings

	Profiles	Postings
Mean # Skills	13.707	13.092
SD	15.14	10.58
Min. # Skills	1	1
Max. # Skills	517	450
N	131,454,168	332,794,806

The Emsi data is comparable to Burning Glass Technologies (BGT) data used in a number of relatively recent papers but has the advantage of containing the much richer skills information described above.[2] Papers using the BGT data often state that it is close to the whole of job-postings. For instance, Burke et al. (2019) claims that the 159 million postings from the Burning Glass data between 2007-2017 constitute a "near-universe". Meanwhile, the Emsi data contains more than 168 million postings between 2010, the earliest year in the data, and 2017. This indicates that the coverage of the Emsi postings is at least as good as the breadth of the Burning Glass postings.

However, the use of a novel dataset does present challenges. First, it is not clear to what extent online job postings are representative of the distribution of occupations across the United States. Atalay, Sotelo, and Tannenbaum (2021), the only other work using the Emsi data of which we are aware, conducts a validation of the Emsi posting data. They find that the number of postings per industry largely mirror opening rates shown in the BLS's Job Openings and Labor Turnover Survey (JOLTS), and the educational requirements in a commuting zone's postings are strongly correlated with the average education of that zone's workers.

[2] For an example of a paper using the Burning Glass data, see Deming and Kahn (2017), "Skill Requirements across Firms and Labor Markets" in NBER's working papers.

There are two main advantages of identifying skills directly from text in job postings and resumes over using proxies such as educational attainment or occupation. The first is that the definitions of skills are highly granular, allowing for heterogeneity of skills even within narrowly defined occupational categories. The aforementioned work by Atalay et al. (2021) documents substantial heterogeneity in occupational task requirements across locations using only the Emsi postings data. Our work is complementary, but distinct, in that we focus on skills rather than tasks in job postings *and resumes*. Skills are anything that defines someone's knowledge, experience, and abilities. They are worker-oriented and common to both workers and jobs. Tasks, on the other hand, are job-oriented activities. A skill may be applicable to many tasks and vice versa.

Table 2 lists the Emsi parsed skills and the top ten O*NET skills for the associated 5-Digit Standard Occupational Classification Code (SOC) for a randomly selected job posting in our data. The job posting is for a "Rooms Inspector" in the 6-digit NAICS industry of Hotels and Motels. Rooms Inspectors fall within the 5-digit SOC 51-1011 "First Line Supervision of Production and Operating Workers". Emsi parsed only 3 skills in the posting for Hotel and Motel Rooms Inspector: Quality Control, Cleanliness, and Training and Development. These three skills do not appear in the top 10 most important occupational skills for SOC 51-1011. Cleanliness doesn't appear as a skill at all and quality control receives an importance rating of 50. While this is only one job posting, it demonstrates the substantial advantage to characterizing a job's skill requirements using the parsed text relative to using occupational skill requirements. The second advantage to using the parsed skills text from posting and profiles is that, unlike other work matching publicly available occupational employment data to O*NET occupational skills data, the skills in postings and profiles do not necessarily represent an equilibrium outcome.

Table 2. Room Inspector Skill Comparison

Emsi Skills	O*Net SOC 5-digit Skills	O*Net Importance
Quality Control	Active Listening	72
Cleanliness	Speaking	72
Training and Development	Time Management	72
	Management of Personnel Resources	72
	Critical Thinking	69
	Monitoring	69
	Social Perceptiveness	69
	Coordination	69
	Judgment and Decision Making	66
	Reading Comprehension	63
	Writing	53

1.4. Measuring Skill Specificity

Given the lack of a consensus method for characterizing “high-skilled” workers and/or jobs, it is not surprising that there is not a consensus for the characterization of distinct skills. Instead of imposing an arbitrary classification of skills, we develop a measure of “skill specificity” that takes advantage of the rich skill information in our data to characterize skills based upon their relationships with other skills in the data. Intuitively, the idea is that a skill is less “specific” if it is closely related to many other skills in the data.

Our measure of skill specificity utilizes a novel data processing method in which the co-occurrence of skills in documents organically creates "skill networks" that map the relationships among skills. Networks are composed of vertices and edges. For our purposes, the vertices are skills listed in resumes. Edges describe the relationship between the vertices (skills) in the network. This relationship can be symmetrical, which is generally referred to as an undirected network, or asymmetrical, which is generally referred to as a directed network. We use a directed network in which edges are weighted by a conditional probability discussed below.

A skill is called "specific" if it has a lower centrality in the network. Broadly speaking, network centrality encapsulates information about a given vertex's position or importance in the network. The primary definition used here is closeness centrality, which is the measure of average weighted distance to all other vertices in the graph. Skills with a low closeness centrality (and thus a high specificity) are, on average, farther away from other skills; when the network is visualized, these skills tend to appear on the periphery of the network.

A key consideration for this method is disentangling a skill being *uncommon* with a skill being *specific*. For example, the skill of being able to use Microsoft Word should have similar specificity to comparable, but less common, word processing software (LibreOffice, for example). An undirected network, where edges are weighted by the probability of co-occurrence in a document, is not well suited to satisfying this requirement. More common skills will tend to have higher edge weights regardless of their actual relationship with the other skill.

The use of asymmetric measures helps separate the notions of commonness and specificity. This work borrows from the field of association rule learning and market-basket analysis (MBA), wherein a researcher desires to find rules governing the co-occurrence of items together in a bundle. The canonical setting is the determination of relationships between goods in a supermarket based on items appearing together in a cart; for instance, $Cream\ Cheese \rightarrow Bagel$, and $\{Peanut\ Butter, Jelly\} \rightarrow Bread$. In our setting, the "basket" is a job posting or a profile, while the "item" is a skill, but considering many-to-one association rules is outside the scope of this paper. A core focus of the field of association rule learning is how to quantify the interestingness of a relationship. We use two such measures here: confidence and conviction (Tsur et al., 1997). Confidence is simply a measure of the conditional probability that skill s_j appears in a document given that skill s_i is observed. The conviction, meanwhile, is the ratio of the probabilities of skill s_j not appearing in a document (without knowledge of s_i) and the probability of skill s_j not appearing when we observe s_i. The two measures are shown in equations 1a and 1b.

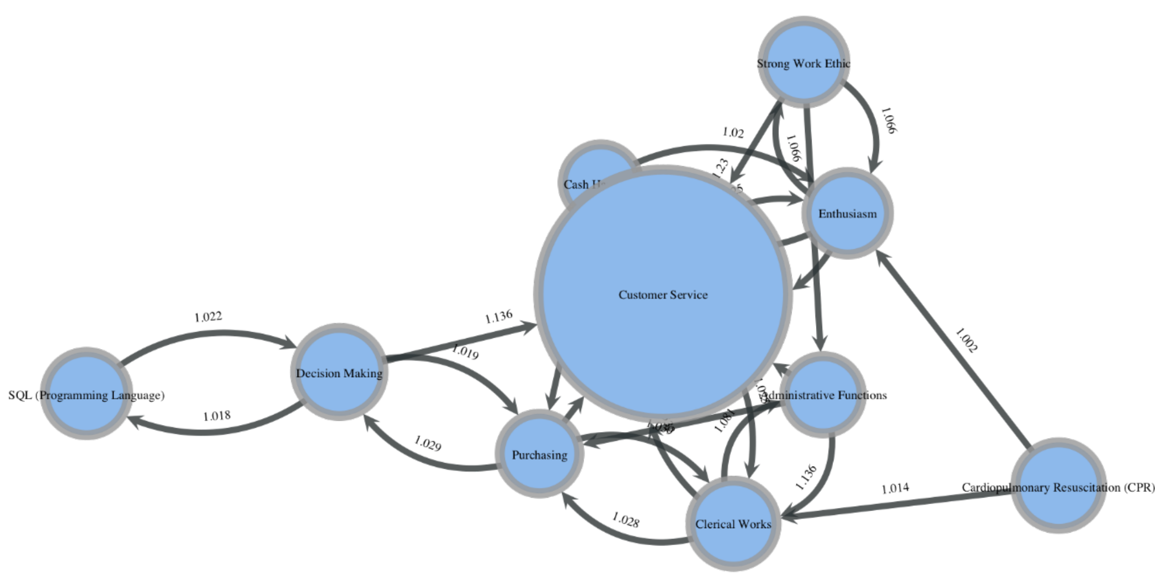
$$(1a):\quad Confidence(s_i \rightarrow s_j) = \frac{Support(s_i, s_j)}{Support(s_i)} = P(s_j|s_i)$$

$$(1b):\quad Conviction(s_i \rightarrow s_j) = \frac{1 - P(s_j)}{1 - P(s_j|s_i)}$$

Intuitively, conviction gives a sense of the value-added of having knowledge about s_i when predicting s_j. The usefulness of this measure is evident when considering skills that are ubiquitous. As an example, the most common skill in the data is *communications*, appearing in roughly 33% of job postings. Even if listing *communications* were truly unrelated with listing another particular skill, the confidence (conditional probability) of any relationship with *communications* as the consequent would be expected to be about 0.33. Although such relationships are nominally strong, knowing the antecedent skill in this case would not change our ability to predict the presence of *communications*. When constructing the network, we use conviction as a parameter to filter edges by their predictive capability.

Figure 1 displays a simple network of ten skills where only edges with conviction greater than one are kept. Skills closer to the periphery of the network, like SQL in Figure 1, will have lower (closeness) centrality, and thus higher specificity. The reader should note that the numbers next to the edges are conviction, whereas we actually weight edges using confidence. For visual clarity, nodes in this graph have at most three outbound edges, but in execution this is not a limitation we make.

Figure 1. Example Network with Ten Skills



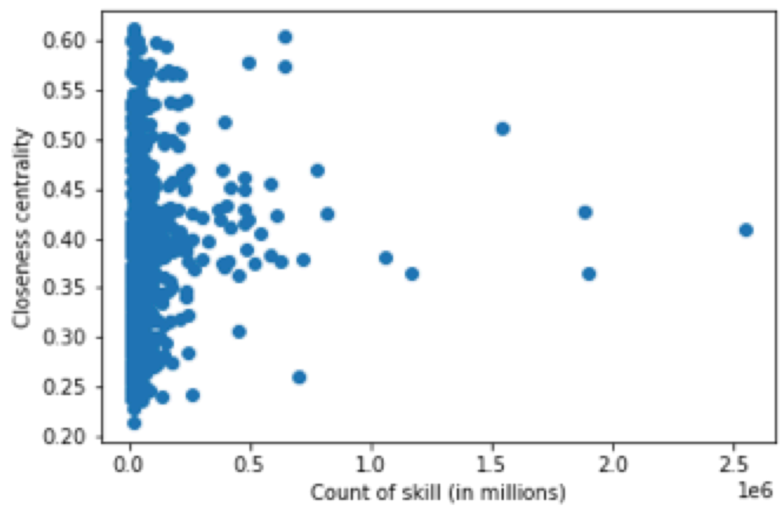
Importantly, the network method allows specificity to be defined at the skill level rather than at the occupation, job, or firm-market level. Previous work uses notions of scarcity or specialization that compare the attribute (e.g., tasks, etc.) vectors between two jobs or occupations using methods to characterize the degree of (non-) overlap, e.g. cosine dissimilarity. The overlap measures have then been aggregated to describe the degree of specialization at the job, occupation, or firm-market level. Instead, our new measure uses information on the network of relationships within a set of documents to define the specificity at the skill level.

In practice, we calculate conviction and confidence for each skill using resumes and postings. We then limit the skill relationships considered in the directed network to only those with a conviction measure greater than 1. In other words, we only consider skill pairs in which knowledge that one of the skills appears in the resume or job-posting improves our ability to predict the appearance of the other skill. We then generate a directed network in which the edge weights are given by confidence and calculate the closeness centrality for each skill in the network. This is our measure of skill specificity.

Figure 2 depicts the relationship between our skill specificity measure and the raw count of resumes/job postings in which the skill appears. Figure 2 indicates no discernable relationship between centrality and counts, suggesting our network approach successfully distinguishes between skills that are "uncommon" and skills that are "specific".

To give a better sense of which skills might be found at each end of the specificity distribution, Appendix Table A1 lists the five most and least specific skills according to our measure. Skills like "civil engineering", "microbiology", and (treatment of) "substance abuse" are considered specific because they have low centrality. They are, on average, the farthest from other skills in the network. Meanwhile, skills like "presentations" and "sales forecasting" are the most central and therefore the least specific. In general, skills that appear with a wide variety of other skills will tend to have a lower specificity. Therefore, it should be noted that some skills may be ‘skilled’ in the traditional sense - i.e., require a significant investment of time to learn - but still be considered less specific.

Figure 2. Relationship Between Closeness Centrality and Count of Skills



1.5. Skill Specificity Across Labor Markets

To better understand skill specificity across labor markets, we recalculate skill specificity measures by labor market. We create a separate network for each labor market using the maximum of a random sample of 3000 skills or the total number of skills in the profiles/resumes from each Commuting Zone. We limit the number of skills in the network to ease comparison of the distribution of specificity within CZs that might be mechanically driven by increased labor market thickness rather than differences in the underlying distribution of skills. We then analyze the labor market specificity within each network and compare. The results indicate skill specificity is more concentrated in larger labor markets.

As an example, Figure 3 plots the labor-market specific distribution of skill specificity for the New York City CZ (134), CZ 270, and CZ 147. CZ 270 is the commuting zone at the 75 percentile of the 2010 population distribution and Cochran, Crosby, Garza, Hockley, Lubbock, Lynn, and Terry counties in Texas. CZ 147 is the 50th percentile population commuting zone in 2010 and includes Camden, Chowan, Currituck, Dare, Pasquotank, and Perquimans counties in North Carolina. There is a clear rank ordering of the mass of specificity by CZ population, with New York having substantially more specific (lower closeness centrality) skills in job postings than the lower population CZs and the smallest CZ having with more mass in less specific (higher closeness centrality) skills.

Figure 3. Job Posting Skill Specificity Distributions in the 99th, 75th, and 50th population Percentile Commuting Zones

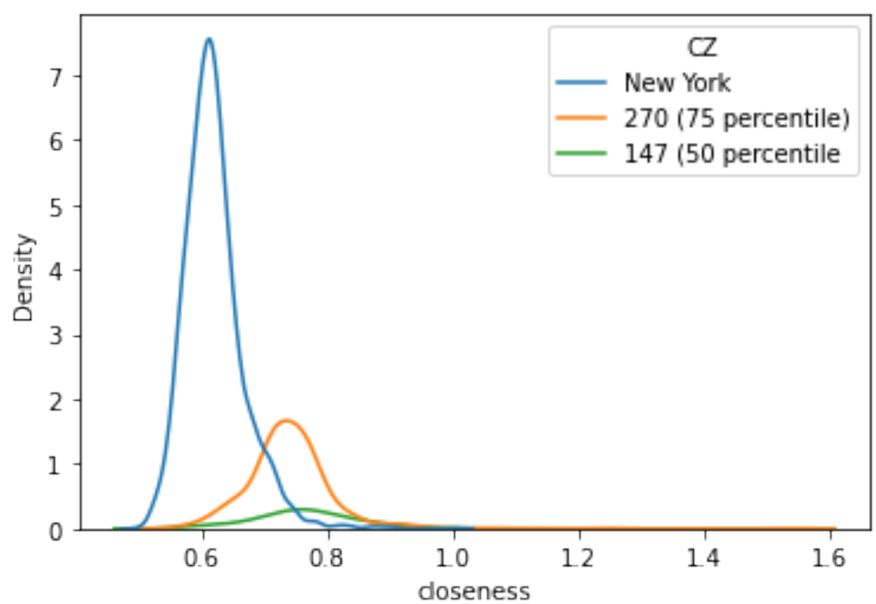


Figure 4 provides a more systematic summary of the relationship between specificity and labor market size. Using the labor market specific networks, we calculate the shortest path between each skill pair in the network and take the maximum value as our measure of network size. In other words, we calculate the confidence-weighted distance between the least closely related skills in the network. This gives us a measure of the size of the network and the extent to which skills in the network are more specific. Figure 4 depicts the relationship between network size and (log) CZ population. It demonstrates that, in general, the skills demanded within a labor market are less closely related, i.e. more specific, as population increases.

Figure 4 does suggest that the relationship between the demand for specific skills and labor market size might be different for very large commuting zones. Thus, in Figure 5, we split commuting zones into two groups at the 75th percentile of 2010 population, with panel A containing the relationship for CZs with populations below the 75th percentile. The figure reveals that skill specificity is increasing in population for both groups but at different rates.

Figure 4. Relationship Between Labor Market Skill Network Size and Population

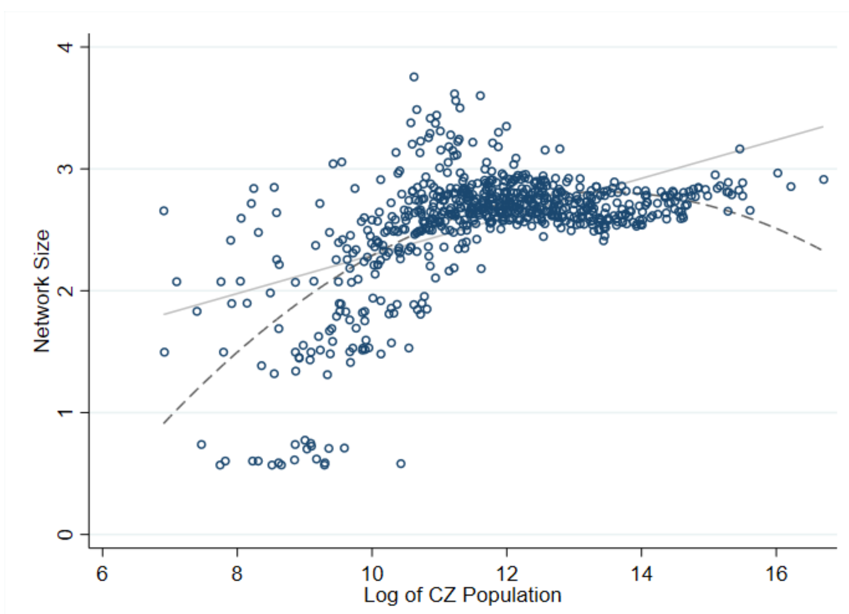
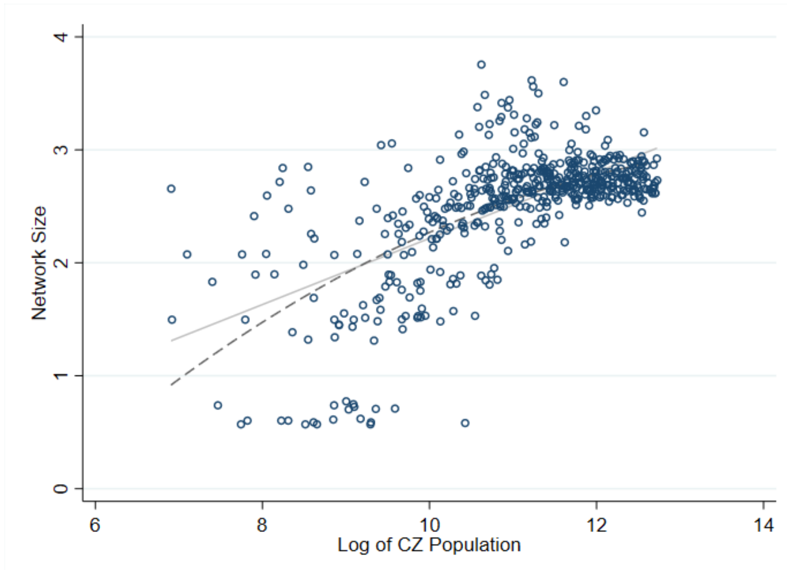
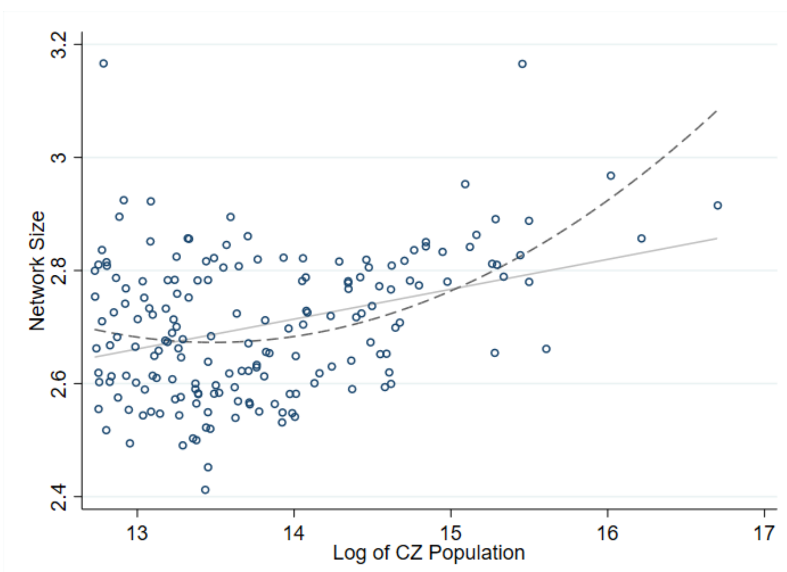


Figure 5. Relationship Between Labor Market Skill Network Size and Population for CZs, Above and Below 75th Percentile of Population

(a) below 75th percentile



(b) above 75th percentile



1.6. Measuring Labor Market Match Quality

We measure labor market match quality at two levels: the (commuting zone, skill) level and the commuting zone level.

1.6.1 Commuting Zones/Skill: Skill-to-Skill Match Quality

To measure mismatch by skill and commuting zone, we simply take the absolute value of the difference in the proportion of resumes listing a given skill and the proportion of job postings listing that same skill. In words, the assumption underlying this metric is that the market for a skill in a commuting zone is better matched when that skill appears in job postings and resumes at a similar rate. This measure of mismatch considers only the distributions of skills demanded and supplied in the market, and waves away market frictions or search that may lead to lower quality matches.

1.6.2 Commuting Zones: Simulated Labor Markets

As an alternative to the simple measure mentioned above, in this section we simulate a labor market in which job postings and resumes are matched together based on the similarity of the skills each contains. Similarity is determined through word embeddings. We restrict to job postings created in 2022 and resumes which were last updated in 2022. In addition, we further constrain the sample to only those documents which contain between 8 and 16 skills. These values approximate the mean number of skills in documents in 2022 plus/minus half of its standard deviation. In total, the matching routine encompasses approximately 95,420,826 skills in 8,118,397 job postings and 104,771,026 skills in 8,897,082 resumes.

Limiting to documents with a certain number of skills listed provides two main benefits. First, it ensures that documents with wildly different numbers of skills are not compared. Documents can have as few as zero skills extracted from them and as many as hundreds. While the algorithm shown does normalize mismatch based on the number of skills listed in documents, this restriction further eliminates noise originating from document pairs with large differences in skill counts. Second, the limitation makes the matching algorithm computationally tractable by limiting the sample size.

1.6.2.1 Word Embeddings. Word embedding is a method from natural language processing that models the semantic meaning of words in a continuous, numerical vector space. Models of this class take a vocabulary of V tokens (words) to be encoded from a document in which those words can appear together in a meaningful context; e.g., a sentence. An embedding model uses the context in which words appear to generate a vector of $W < V$ weights for each word. Because words which are closer in meaning will also be represented by closer vectors, they are an attractive method for measuring the similarity and differences between the meanings of words. A commonly used example illustrates the effectiveness of this technique. Suppose the model has embedded the words "king", "queen", "man", and "woman". Then, the operation on their vectors $king - man + woman \approx queen$ will hold true given sufficient training.

The fact that embeddings allow computation of the difference between two given words is useful for our purpose. To better understand this, consider the following: are the skills "Stata" and "R" more or less similar than "C++" and "Marketing"? Human intuition permits a judgment on this proposition, but, prior to this work, it was unclear how to quantify these comparisons. The use of embeddings allows us to flexibly estimate such similarities and differences. Rather than embedding words, we embed the skills extracted by Emsi (which may be groups of words). Likewise, the context from which the model learns the meaning of a token comes not from co-occurring in a sentence but from co-occurring in a job posting.

Multiple methods for generating word embeddings exist. Such methods include GloVe[3], an approach which conducts dimensionality reduction on the word (skill) co-occurrence matrix, and the more popular word2vec[4], a neural-network based method which predicts words (skills) based on the words (skills) which appear in a given-sized window around them in context. We use word2vec for primarily logistical reasons. Word2vec is relatively fast computationally, uses less memory compared to GloVe and other exhaustive models, and has a readily available and well-documented implementation[5].

We train word2vec on skills appearing in job postings between January 2018 and October 2022. We limit training to the 10,000 most frequently appearing skills, use a window size of five, and train the model over five epochs. This results in a vector, for each skill, of 100 numbers. To measure the distance between skill embeddings, we use a weighted function of the cosine similarity between the two embedding vectors. Appendix Table A2 compares the ten most similar skills found to Stata when using both cosine similarity and the Euclidean distance metrics. While the two yield different sets of results, both sets seem plausible, so the use of cosine similarity is ultimately author preference.

1.6.2.2 Labor Market Simulation. We use the word embeddings to determine the similarity between the skills in job postings and resumes. The algorithm is as follows. Let the i^{th} job posting be represented by the vector $\vec{\mathrm{j_i}} = (\mathrm{j}_{\mathrm{i}1}, \ldots, \mathrm{j}_{\mathrm{ik}}, \ldots, \mathrm{j}_{\mathrm{is}})$, where $\mathrm{j}_{\mathrm{ik}} = 1$ if skill k appears in posting i. Likewise, let the j^{th} resume be represented by the vector $\vec{\mathrm{r_j}} = (\mathrm{r}_{\mathrm{j}1}, \ldots, \mathrm{r}_{\mathrm{jk}}, \ldots, \mathrm{r}_{\mathrm{js}})$, where $\mathrm{r}_{\mathrm{jk}} = 1$ if skill s appears in resume j. The distance between a given posting and resume is

[3] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation.
[4] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space
[5] https://radimrehurek.com/gensim/models/word2vec.html

expressed by the following expression, where D is an $s \times s$ matrix where the k, lth element of $D_{kl} = \text{distance}(\text{skill}_k, \text{skill}_l)$. The terms $\sigma(j_k)$ and $\sigma(r_l)$ denote the total number of skills listed in j_k and r_l respectively. We divide by these terms in an attempt to prevent the number of skills listed from having an effect on the matching process.[6]

$$(2)\ distance(j_k, r_l) = \frac{\overrightarrow{j_k} \cdot D \cdot \overrightarrow{r_l}' \cdot}{\sigma(j_k) \quad \sigma(r_l)}$$

We measure the distance between two skills as the distance between their embedding vectors, as discussed in a previous section. In this version of this report, we use the cosine similarity between the two vectors instead of the Euclidean distance. The cosine similarity yields values in the interval [-1, 1], so we transform them to be in the interval [0, 1] as follows. The tuning parameter α controls the preference for skill similarity in the matching process. For our analysis, we set $\alpha = 4$.

$$(3)\ distance(skill_k, skill_l) = (1 - \frac{similarity(skill_k, skill_l) + 1}{2})^{\alpha}, \alpha > 1$$

For every commuting zone, each job posting is matched to the resume that minimizes the distance metric above and an average of mismatch across jobs is calculated. Appendix Table A3 shows several example matches. Matches are mutually exclusive. Once a resume has been chosen for a given job, it is not eligible for being matched with another job. Because of this, the order in which the jobs are matched will partially determine, on average, the effectiveness of individual matches. In short, jobs that are matched last will have a smaller pool of possible

[6] To be clear, this is because adding a skill to either the job or resume vectors will mechanically increase the product given that all elements are nonnegative.

resumes to choose from, which may lead to a worse match compared to if it were matched first. To mitigate this, we conduct 30 trials of this matching routine where the order of postings is randomized each time.

The method we use to measure the distance between two groups of embeddings could be improved upon[7]. For instance, the Word Mover's Distance (WMD) matches two documents by first finding closest matches for individual word embeddings between documents, and then calculating the overall distance metric from the distances of individual pairs. However, due to the size of our data, computational limits create an impediment for such improvements. The benefit of our current approach is that Equation 2, our mismatch equation above, can be "stacked" to create the matrix multiplication shown below in Equation 4. Matrix multiplication is already highly efficient in standard statistical software, and hardware acceleration through the GPU can provide orders of magnitudes of further speed improvement. Below, J is an $n \times s$ matrix wherein each row is a sequence of indicators for a given skill being requested by a job. Likewise, R is an $m \times s$ matrix wherein each row, corresponding to a resume, is a vector of indicators for a given skill being provided. N is a matrix with $N_{i,i}$ counting the number of skills in the i^{th} job posting, while M is the equivalent for resumes.

$$(4)\ A = N^{-1}JDR'M^{-1}$$

Because each trial in the simulation yields the average mismatch between job posting / resume pairs in the commuting zone, the end result of the Monte Carlo matching routine is, for

[7] Another alternative would be to use Doc2Vec, which can create embedding vectors from lists of tokens (skills, in this case), or to simply compare the means of the word embeddings.

each commuting zone, a distribution of average match quality across 30 trials[8]. From these trials, an average (of averages) of posting-resume mismatch is generated for each commuting zone.

1.7. City Size and Match Quality

1.7.1. Empirical Specification

We conduct a regression analysis for each mismatch metric. In short, for the measure of mismatch at the level of commuting zone - skill, we regress the mismatch for each skill in the CZ on the log population of that CZ in 2021. For the metric generated by the simulated labor market, average skill mismatch by CZ is regressed on log of population. Because our main question is whether urban growth promotes better matching, the coefficient of interest for both regressions is on the log of population.

Identification of the marginal effect of urban density on match quality using OLS requires that the random component of match quality be uncorrelated with urban density. This requirement could be broken if a missing regressor is correlated with density and influences match quality; for instance, the total factor productivity of an urban area may be correlated with both density and match quality. Likewise, violation could occur if density and match quality are determined simultaneously. If people are attracted to an urban area because of a labor market that generates good matches, the estimated marginal effect will overstate the effect of density on match quality.

Ciccone and Hall (1996) demonstrate that this endogeneity issue can be surmounted by using a set of instrumental variables which affect density but are plausibly uncorrelated with the error term. Following best practices outlined in Combes and Gobillon (2015), we estimate the effect of population on match quality using 2SLS using a deep lag of population density, the

[8] In practice, the variance of this distribution is extremely small. This might not be the case if the data had fewer observations.

1920 population, as an instrument for modern density of each CZ. The validity of the deep lag of population as an instrument requires that the spatial distribution of population in the past explains the current spatial distribution of population and that the local drivers of worker match quality in the past differ from those of today (Combes, Duranton, Gobillon, and Roux, 2010). While the population density of an area in 1920 is correlated with a city's current density, we assume it will not affect match quality except through its relationship with current density. If this assumption holds, our estimator will identify the causal effect of density on match quality. To gauge bias which might arise from this issue, we present estimates using both OLS and 2SLS.

As a robustness test, we also present estimates using several other common instruments: a deeper lag of population, a seismic risk index, landslide risk index, and an indicator for sedimentary bedrock presence. The logic of using geological instruments rests on the notion that geological factors help explain the spatial distribution of population. For instance, areas with sedimentary bedrock can support higher population densities, and areas with geological instability provide risks that may have deterred settlement. Identification using these instruments requires that geology no longer be a relevant factor in determining match quality except through the way it influenced settlement patterns.

Beginning with the estimation at the commuting zone - skill level, the main model is shown in Equation 5a. The outcome, $mismatch_{cs}$, is the absolute value of the difference in the percentage of resumes and job postings in which the given skill *s* appears in commuting zone *c*. On the right hand side, $log(pop_{c,2021})$ is the log of the commuting zone's population in 2021, while $1(high_s)$ and $1(low_s)$ are indicators for skill *s* being in the highest and lowest tertile of specificity, respectively, where skills with higher closeness centrality are called less specific. In other words, $1(high_s)$ can also be considered an indicator for being in the lowest tertile of closeness centrality.

As discussed above, we use 2SLS to correct for the endogeneity of 2021 population with mismatch. Because we interact the endogenous regressor with the high and low specificity indicators as we do in 5a, identification requires that we instrument for the resulting endogeneity of the interaction terms. For an endogenous interacted term, the interaction of the instrument and the indicator is a sufficient new instrument (Wooldridge, 2001). Therefore, Equations 5b, 5c, and 5d present the three first-stage equations that allow the model to be just identified. In each of these equations, the endogenous term is regressed on $pop_{c,1920}$, the population of the commuting zone one-hundred years prior in 1920, as well as interactions between the deep lag of population and the high and low specificity indicators. When estimating Equation 5a via 2SLS, we use estimated values of the log of population and the interaction terms in 5a from the first-stage equations.

Estimation of Equation 5a will yield the marginal effect β_1 of urban population on skill-level mismatch. A negative value for β_1 indicates that urban growth decreases the mismatch for skills on average; in other words, a negative value for β_1 means that urban growth increases match quality. Our secondary research question is concerned with whether skills experience differential matching along the distribution of specificity, and this question is answered through the coefficients β_2, β_3 on the interaction terms. The coefficient on β_2 (β_3) measures the change in matching premium for high (low) specificity skills.

$$
\begin{aligned}
(5a)\ mismatch_{cs} &\\
&= \beta_0 + \beta_1 \log\left(\widehat{pop_{c,2021}}\right) + \beta_2 1(high_s) \times log\ (\widehat{pop_{c,2021}}) \\
&+ \beta_3 1(low_s) \times log\ (\widehat{pop_{c,2021}}) \ + \epsilon_{cs}
\end{aligned}
$$

$$(5b): \log(pop_{c,2021}) \\ = \alpha_0 pop_{c,1920} + \alpha_1 1(high_s) \times pop_{c,1920} + \alpha_2 1(low_s) \times pop_{c,1920} + \xi_{cs,1}$$

$$(5c): 1(high_s) \times \log(pop_{c,2021}) \\ = \gamma_0 pop_{c,1920} + \gamma_1 1(high_s) \times pop_{c,1920} + \gamma_2 1(low_s) \times pop_{c,1920} + \xi_{cs,2}$$

$$(5d): 1(low_s) \times log\,(pop_{c,2021}) \\ = \delta_0 pop_{c,1920} + \delta_1 1(high_s) \times pop_{c,1920} + \delta_2 1(low_s) \times pop_{c,1920} + \xi_{cs,3}$$

The second half of our analysis focuses on mismatch at the commuting zone level. We conduct a routine in which job postings are matched to resumes based on the skills contained within each. This method uses embeddings of the skills to glean context based on their co-occurrence, a process which is discussed in greater detail in Section 6.2.2 above. Equation 6a and 6b represent the second and first stages respectively, where $mismatch_c$ is the mean of the distribution of average mismatch from 30 trials. Lower values of $mismatch_c$ represent better matches, on average, between job postings and resumes in commuting zone *c*.

$$(6a)\ mismatch_c = \beta_0 + \beta_1 log\,(\widehat{pop_c}_{,2021}) + \epsilon_c$$

$$(6b)\ log\,(pop_{c,2021}) = \alpha_0 + \alpha_1 pop_{c,1920} + \xi_c$$

As cities grow, a portion of the decrease in the average mismatch of the best-matched worker-firm pairs is strictly mechanical. To illustrate this, imagine a labor market where workers pair with a job that has a mutual best-match given available options. The labor market begins with a single job filled by a single worker. A new job opens, leading the worker to reevaluate and make a choice to stay or switch jobs. Because only match-improving pairings are made, then all else equal a new worker or job can only decrease the mismatch of pairs containing existing workers or jobs. On the other hand, new pairings composed only of entrant or replaced workers and jobs may have a worse match than average, pulling the mean mismatch down. In practice, this latter scenario seems unlikely to outweigh the magnitude of the former effect.

We will refer to this mechanical effect as the ‘market thickness’ effect. However, other drivers of the matching premium exist. Dauth et al. (2022) distinguishes two types of assortative matching that are relevant to our analysis: between-city and within-city assortative matching. Between-city assortative matching occurs when workers sort to different cities based on their expectation of the labor market, for instance in the case of high-quality workers sorting to cities with high-quality firms. On the other hand, within-city assortative matching relates to how the given workers and firms in a city match. While the role of market thickness in the urban contribution to match quality is important[9], we seek to capture these other policy-relevant sources of the urban matching premium.

To do so, we estimate two additional variants of Equation(s) 6. The first variant randomly samples 10,000 resumes and 10,000 job postings from each commuting zone before beginning matching. Using the same number of documents in each labor market eliminates the possibility that results are driven merely by the mechanical process described above. Further, comparing the results between the unrestricted and restricted samples gives us an idea of the relative magnitudes of the mechanical and non-mechanical drivers of the effect. The second variant restricts the sample again while replacing the matching criterion with random assignment. Rather than the resume that best fits the posting, postings are matched with a random resume.

Our idealized matches forgo frictions faced in a real matching market. Matching based only on skills removes impediments like degree or experience requirements; further, labor

[9] Indeed, Duranton and Puga’s (2004) basic model of urban matching is strictly mechanical.

market agents in the real world do not have perfect information about the entire spectrum of jobs/employees available to them. Therefore, our simulation design implicitly assumes perfect within-city assortative matching. The coefficient of the second variant above can be considered a baseline, but a significant negative coefficient would provide strong evidence of between-city sorting of skills as population grows. For example, one might expect a randomly selected posting and resume to be better matched in a locale like Silicon Valley, a market with a concentrated industry. On the other hand, simulating the labor market as we do in the first variant, with the matching algorithm applied to a random sample, measures both between-city and (perfect) within-city sorting. Therefore, we might expect the true effect of population on mismatch to be bounded between the coefficients on the two variants.

1.7.2. Results

Table 3 reports the results from the ordinary least squares and two-stage least squares analysis using mismatch defined at the commuting zone - skill level for the 3000 skills which appear most frequently in documents. Once again, the outcome is the absolute value of the difference in percentages of resumes versus job postings in which the skill appears. Starting with the main OLS specification in Column 1, the table shows that the OLS coefficient for the log of population is negative and significant, indicating that a 1% increase in population decreases skill-level mismatch by .00122 on average. The interaction between the high specificity indicator and the log of population is negative and significant, while the interaction between the low specificity indicator and the log of population is positive and significant. Taken together, this means that skills with greater specificity see a larger reduction in mismatch as population increases.

Table 3. Effect of Log of Population on Mismatch of Top 3000 skills at the Commuting Zone - Skill Level, with Interactions for Skill Specificity

	OLS		IV	
	1	2	3	4
log(pop$_{c,2021}$)	-0.109***	-0.328***	-0.070***	-0.194***
	(0.0008)	(0.0016)	(0.0014)	(0.0030)
1(high$_s$) × log(pop$_{c,2021}$)	-0.004***	-0.005***	-0.005***	-0.007***
	(0.0002)	(0.0004)	(0.0002)	(0.0005)
1(low$_s$) × log(pop$_{c,2021}$)	0.048***	0.076***	0.050***	0.082***
	(0.0002)	(0.0004)	(0.0002)	(0.0004)
Constant	1.55***	4.59***	1.06***	2.86***
	(0.0097)	(0.0210)	(0.0178)	
Job requires degree?	No	Yes	No	Yes
N	1,325,622	1,067,861	1,309,156	1,055,067

The primary 2SLS specification is reported in Column 3. The coefficient on the log of population is once again negative and significant, but smaller in magnitude than in the OLS specification. The signs on the interaction terms are consistent with the signs of the respective OLS coefficients but have larger magnitudes. Regarding the first stage, our inclusion of terms interacted with both the endogenous regressor and the instrument results in artificially high F statistics which are not useful to report. Instead, Appendix Table A4 reports, for a variety of instruments, the first-stage F when Equation 5 is estimated with no interactions on the endogenous regressors or the instruments. The population in 1920, with an F of 1529.47, is a highly relevant instrument.

We also estimate this model amongst only job postings requiring at least a bachelor's degree for better comparison to other work that often proxies skill with education. Column 2 shows the results when the model is estimated using OLS on the restricted sample. The coefficient on the log of population decreases to -0.348 after restricting to degree-requiring postings, approximately tripling in magnitude. Likewise, the coefficients on the interaction terms also increase in magnitude after the sample restriction. When the 2SLS model is estimated with this restriction, as in Column 4, the pattern of results is similar to that of the OLS estimation. Once again, the coefficient on the log of population nearly triples, with the magnitude of the coefficients on the interaction terms also growing.

As a robustness test, we also estimate Equation 5 with a deeper lag of population – 1870 rather than 1920 – and several other standard geological instruments. Appendix Table A5 reports these results. In summary, our results are mostly insensitive to instrument choice. The coefficient on log of population is negative and significant for all instruments but sedimentary bedrock presence, with values ranging from -0.054 to -0.109. However, instrumenting with sedimentary bedrock presence, the least relevant instrument according to Appendix Table A4, causes the sign to flip and yields a positive and significant coefficient of 0.069. The signs of the interaction terms remain consistent and are precisely estimated for all choices of instruments.

Next, we repeat the analysis shown in Table 3 while restricting to only the 1,000 skills most frequently appearing in documents. By doing this, we test whether the patterns observed in Table 3 hold among a smaller subset of more common skills. To the extent that there are doubts regarding the distinction of a skill being less specific (more general) versus it being common, comparison of patterns of results among more common skills may help address them. After restricting to the 1,000 most frequent skills, the high and low indicators now refer to a skill being in the highest specificity tertile of the 1,000 skills, rather than the previous 3,000.

Table 4 demonstrates these results. Overall, the pattern of results is nearly identical to those shown in Table 3. For the coefficient on the log of population, the OLS estimate among degree-requiring postings is the largest, followed by the 2SLS estimate among degree-requiring postings, then the unrestricted OLS estimate, and finally the unrestricted 2SLS estimate. However, compared to Table 3, Table 4 shows greater heterogeneity in the matching premium by high and low specificity.

Table 4. Effect of Log of Population on Mismatch of Top 1000 Skills at the Commuting Zone - Skill Level, with Interactions for Skill Specificity

	OLS		IV	
	1	2	3	4
$log(pop_{c,2021})$	-0.137***	-0.382***	-0.073***	-0.211***
	(0.0018)	(0.0034)	(0.0034)	(0.0066)
$1(high_s) \times log(pop_{c,2021})$	-0.005***	-0.024***	-0.010***	-0.012***
	(0.0006)	(0.0011)	(0.0011)	(0.0020)
$1(low_s) \times log(pop_{c,2021})$	0.082***	0.163***	0.069***	0.118***
	(0.0006)	(0.0012)	(0.0011)	(0.0020)
Job requires degree?	No	Yes	No	Yes
N	559,892	477,604	552,080	471,412

Next, Table 5 presents the results from the estimation of the effect of population on mismatch at the commuting zone level generated by algorithmically matching job postings and resumes across commuting zones. It encompasses resumes and job postings created or updated in 2022 which had between 8 and 16 skills extracted. Three specifications are estimated for both OLS and 2SLS. The first, basic, variant matches all in-sample resumes and postings using a matching criterion. The second restricts the sample to 10,000 randomly selected postings and 10,000 randomly selected resumes per labor market but is otherwise the same. The third also randomly samples but instead randomly matches postings with resumes. For comparability between the three variants, the sample is further restricted to commuting zones where the random sample binds – in other words, those with at least 10,000 resumes and 10,000 postings meeting our criteria in 2022. The outcome here is the mean of the distribution of average minimum mismatch across 30 trials for each commuting zone, converted to a z-score.

When estimated using OLS and the first simulation variant, depicted in column (1) of Table 5, the coefficient on the log of current population is negative and significant (-0.394); from this result, we can infer that a 1% increase in population decreases mismatch between job postings and resumes by about .004 standard deviations on average. Likewise, doubling population would decrease mismatch by about 0.27 standard deviations on average[10]. Figure 6 below visually depicts the clear negative relationship between urban population and mismatch. When the endogeneity of current population is corrected by instrumenting using the population in 1920, depicted in column (4), the coefficient on the log of population grows to -0.693 and remains significant, indicating downward bias of the OLS coefficient.

When we instead randomly sample 10,000 job postings and 10,000 resumes from each labor market before matching, the resulting coefficient attenuates greatly. For both the OLS and 2SLS specification, the random-sampling model coefficient is about 20% of the size of the coefficient in the standard model. This fact implies the "market thickness" effect described earlier plays a larger role than either between or within-city assortative matching. When the matching criterion is removed in columns (3) and (6), the coefficient goes to zero. Compared to smaller cities, randomly selected resumes in large cities do not seem to be more closely matched to randomly selected job postings.

[10] Effect of a 100% increase: -0.394 * ln(2) = -.273

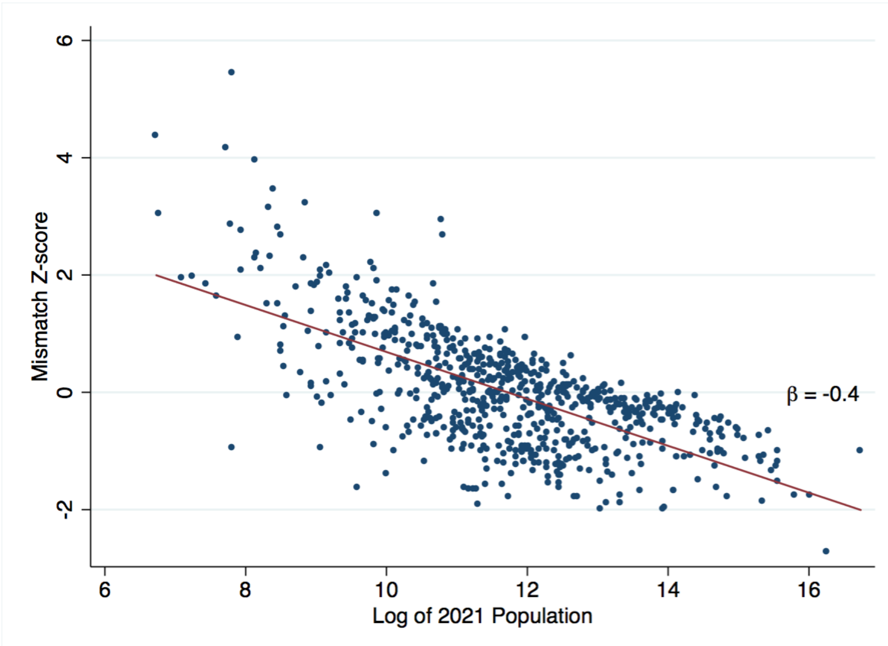
Table 5. Effect of Log of Population on Commuting Zone Mean Mismatch Z-score from Labor Market Simulation

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of 2021 Pop.	-0.394***	-0.0888***	0.0286	-0.693***	-0.132***	0.0739
	(0.05)	(0.01)	(0.06)	(0.09)	(0.03)	(0.10)
Intercept	4.917***	0.945***	0.0449	9.148***	1.549***	-0.59
	(0.76)	(0.19)	(0.78)	(1.32)	(0.39)	(1.40)
Random sample?	No	Yes	Yes	No	Yes	Yes
Random matching?	No	No	Yes	No	No	Yes
N	128	128	128	126	126	126
R-sq	0.359	0.208	-0.006	0.142	0.157	-0.013
First-stage F:	61.28					

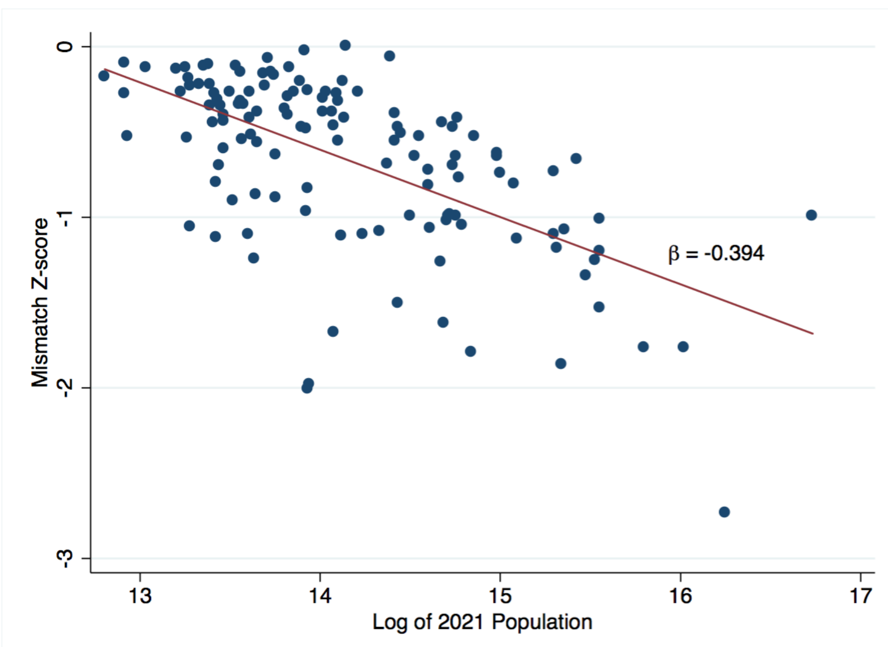
Standard errors in parentheses. *, **, and *** represent significance at p < 0.05, 0.01, and 0.001 respectively.

Figure 6. Relationship Between CZ-level Mismatch and Log of Population

(a) all commuting zones



(b) commuting zones with sufficient resumes and job postings



A finding indicating that randomly chosen documents are better matched in large cities would be strong evidence of between-city assortative matching, yet our findings above do not yield that evidence. However, large cities may contain multiple concentrated industries. In such cases, it is plausible that the skills of the average worker and requirements of the employer would converge within sectors but not in the overall market. Because workers looking for a job - and jobs looking for an employee - are more likely to search within specific occupations (Dauth et al., 2022), we estimate another specification of the model, repeating the procedures described above, while instead matching within two-digit SOC occupational categories. Table 6 depicts these results.

Table 6. Effect of Urban Population on Mean Mismatch Z-score

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of 2021 Pop.	-0.344***	-0.056**	-0.004	-0.462***	-0.0254	0.0413
	(0.031)	(0.017)	(0.037)	(0.048)	(0.029)	(0.076)
Intercept	3.804***	0.104	0.362	5.48***	0.538	0.281
	(0.44)	(0.241)	(0.528)	(0.682)	(0.406)	(1.073)
Random sample?	No	Yes	Yes	No	Yes	Yes
Random matching?	No	No	Yes	No	No	Yes
N	123	123	123	122	122	122
R-sq	0.581	0.081	-0.008	0.522	0.054	-0.017
First-stage F:	59.34					

Standard errors in parentheses. *, **, and *** represent significance at p < 0.05, 0.01, and 0.001 respectively.

The pattern of results broadly follows Table 5 but with generally smaller magnitudes. The matching routine in Table 5 has no search frictions and perfect information, so imposing a restriction – occupational requirement – will naturally attenuate the coefficients. Further, the

2SLS specification no longer detects a statistically significant effect when matches are made within occupations and randomly sampled as in column (5). The coefficient on the log of population remains insignificant in columns (3) and (6). One might expect that, within a city and occupation, a random worker and job posting may be better matched as cities grow; however, our evidence does not support this claim. On the other hand, it is possible that this effect might be detected at more specific levels than the two-digit SOC occupational classifications.

1.8. Conclusion

This research attempts to answer two related questions. The first of these is simple: does increasing the population of cities lead to improvements in match quality between the workers and employers in those cities? Significant effort using a variety of methods has been devoted towards the answering of this question in the past. The answer holds important implications for the economics of urban areas, especially with respect to explaining the microfoundations which drive agglomeration economies, generating large wage and productivity premiums for workers and firms in urban areas. We answer this question using detailed job posting and resume data from the United States, the first time, to our knowledge, this question has been addressed with such data. We measure mismatch at both the commuting zone - skill level and at the commuting zone level, developing a method in which job postings and resumes are matched based on the semantic content of the skills listed within. Our use of word embeddings to do so, taken from the field of natural language processing, may prove useful to others in economics or related fields working on problems trying to derive meaning from item cooccurrence.

Prior research has established that differential gains to productivity from agglomeration economies based on worker skill has driven 80% of the total increase in wage inequality from city size since the 1980s (Baum-Snow, Freedman, & Pavan; 2018). This fact motivates our second question, which asks: does the match premium from urban density for a skill depend on how general or specific the skill is? We ask this in an effort to reframe what it means to be "high-skilled." To answer this question, we employ a novel method for assessing specificity that leverages the implied relationships among skills which cooccur in job postings and resumes. By representing conditional probabilities of skill occurrence as a directed network, we can calculate a centrality measure which we posit is decreasing in their specificity.

Using standard instruments to mitigate for the endogeneity of urban size, we find consistent evidence across all models that urban population increases match quality. From the labor market simulation, we find that doubling city size reduces average mismatch between postings and resumes in that commuting zone by 0.27 standard deviations. Likewise, doubling city size reduces the difference in the percentage of times a skill is listed in job postings versus resumes in a given commuting zone by 0.04 percentage points on average. Further, we conduct labor market simulations to try to disentangle possible sources which could generate the match premium we observe from urban size. Our simulation estimates that roughly 80% of the match premium is explained by the mechanical effect of labor pooling which is unrelated to sorting. More precisely, when sampling the same number of documents in all labor markets, 20% of the original effect remains, indicating the presence of the sorting of workers and firms between cities.

Importantly, we find that skills with high specificity are significantly better matched relative to more general skills as cities grow. The patterns we find do not change when the sample of skills is restricted to only the thousand most frequent ones, indicating a relationship between urban growth and the relative specificity of skills that holds across the distribution of skill commonness. Further, we find that restricting to only postings requiring a college degree amplifies the measured effects, meaning that this phenomenon may be most concentrated in the market for highly educated workers.

Better matching for more specific skills is a plausible mechanism by which the skill-bias of agglomeration economies, which drives a significant portion of nationwide wage inequality, could be explained. Whereas Baum-Snow, Freedman, and Pavan (2018) tentatively suggest the learning microfoundation to be the culprit of skill-biased agglomeration, we have provided initial evidence that the matching microfoundation may be a key driver. Further work is now needed to connect differential matching of skills across the specificity distribution to worker productivity and wage inequality.

CHAPTER II. Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia

(with Tim Sass and Ishtiaque Fazlul)

2.1. Introduction

It is well established that there is a strong link between K-12 performance and later adult outcomes, such as post-secondary education attainment, teenage pregnancy, criminal activity, and adult employment and earnings (Cunha & Heckman, 2007, 2009; Goldhaber & Özek, 2019; Heckman et al., 2010a, 2010b; Heckman et al., 2013; Watts, 2020). Given that differences in educational performance appear early in life and the fact that it is increasingly difficult to remediate disparities in education as children age, many have suggested prioritizing early educational interventions as a means of improving performance both in childhood and later in life (Carneiro & Heckman, 2003; Cunha et al., 2006; Heckman, 2000, 2008). This view has its theoretical foundations in the child psychology literature (Justice et al., 2009; Stipek, 2006) and is supported by early studies of high quality but small-scale Pre-K programs such as the Perry Pre-School Program (Heckman et al., 2010a, 2010b) and the Carolina Abecedarian Project (Garcia et al., 2020), which find substantial benefits to participants in both the short-run and long-run. Fueled in part by evidence from these small-scale experiments like Perry Preschool and Abecedarian, some states initiated or significantly expanded pre-K education programs in the 1980s and 1990s (Mitchel, 2001). While most of these state-funded pre-K programs have income thresholds, as of 2017, 11 state programs (including Georgia's) are "universal" programs that have no income restriction for participation (Friedman-Krauss et al., 2022).

Georgia's Pre-K Program is a state-funded early education program for four-year-old children in Georgia that is administered by Bright from the Start: Georgia Department of Early

Care and Learning (DECAL). The program began in 1993 and its goal is to prepare children for success in Kindergarten and in later school years. Currently, there are approximately 84,000 available slots in Georgia's Pre-K Program spread over roughly 4,000 sites that are located throughout the state (Goldring, 2020). Some programs are located at public elementary schools and are operated by public school districts (school-based pre-K sites, or for the purposes of this paper, SBPK), while others are operated by private child development centers, independent of local school systems (non-school-based pre-K sites, hereafter, non-SBPK).

Currently, little is known about the effects of participating in SBPK programs in Georgia on later educational outcomes. In this paper we estimate the impacts of winning an enrollment lottery and attending a school-based site in Georgia's Pre-K Program on a student's academic achievement, attendance, and discipline in later grades using data from a large school district in the metro Atlanta area (hereafter, the District). Our comparison group are students with similar characteristics who sought admission to an over-subscribed SBPK in the District but did not win the enrollment lottery and did not end up enrolling in any (school-based or non-school-based) site in Georgia's Pre-K Program. Thus, we are not comparing the efficacy of attending a SBPK program relative to a non-SBPK program. Rather, we are comparing outcomes for students in school-based sites in Georgia's Pre-K Program to students whose families sought admission to a SBPK program, but were not granted admission and ended up either attending an early learning program (e.g., a Montessori or private school) outside of Georgia's Pre-K Program or no formal early-learning program at all. This approach enables us to evaluate SBPK against a hypothetical scenario where no GA Pre-K is available, in other words, against a scenario of "business as usual" without GA Pre-K.

In addition to average outcomes, we also show how the effects of enrolling in a SBPK vary based on the sociodemographic characteristics of children, like free or reduced-price meals (FRPM) eligibility, a crude measure of poverty. Finally, we characterize the early childhood education decisions made by families of children who enter lotteries for over-subscribed SBPK sites but do not win the lottery and thus are not offered admission. More specifically, using data from a metro-Atlanta school district, we address the following questions:

1. What is the effect of enrolling in a school-based site in Georgia's Pre-K Program (SBPK) for students who would otherwise not attend Georgia's Pre-K Program on future test scores, attendance, and behavior in K-12?
2. How does the effect of enrolling in a school-based site in Georgia's Pre-K Program (SBPK) vary by families' economic status?
3. How do the enrollment decisions of lottery non-winners vary by a student's demographic subgroup?

Using Pre-K enrollment data from GA Pre-K and admission lottery and roster data from the District, we find that lottery winners enter kindergarten significantly better prepared, scoring around six national percentiles higher on the Measures of Academic Progress (MAP) math and reading tests. However, these gains fade by the end of kindergarten, and some negative effects on achievement emerge by grade 4. The negative effects in later grades may be driven by students in the control group who attend options outside of Georgia's Pre-K Program. We find that free- and reduced-price-meal (FRPM) students benefit more from Pre-K compared to non FRPM students in grades 1, 2 and 4, suggesting that attending pre-K may be more beneficial for disadvantaged students, a common finding in the early education literature (Currie, 2001; Lee et al., 1990). Winners were no less likely than non-winners to commit a disciplinary infraction in any grade. however, they did miss about one fewer day of instruction in each grade after kindergarten. FRPM status does not moderate the effect of Pre-K on attendance and discipline.

1.1 *Background on Georgia's Pre-K Program*

Early education providers in Georgia may apply to become a Georgia's Pre-K Program Provider; upon approval, they receive reimbursement conditional on meeting DECAL guidelines. The level of and requirements for reimbursement are almost identical between the SBPK and non-SBPK sites. For example, conditional on a teacher's level of education and certification, DECAL grants equal funding for teacher salaries at both types of sites, and only slight differences exist between the two in the amount of funding given for non-wage benefits and classroom operating expenses. However, the District studied supplements the DECAL-provided salaries of teachers in school-based sites to match the District's pay scale for K-12 teachers. In short, DECAL guidelines are unlikely to create differences in teacher quality, but differential pay from additional District funding might. It is not clear whether non-school-based sites also supplement teacher funding or the extent to which differences in salary translate to differences in teacher quality. In addition, both SBPK and non-SBPK sites are required to choose from a set of DECAL-approved curricula for instruction. It is doubtful, then, that students in non-SBPK sites will learn significantly different content than those in SBPK sites.

Families whose children are enrolled in either a SBPK or non-SBPK site in Georgia's Pre-K Program face no out-of-pocket costs for regular instruction. Providers in Georgia's Pre-K Program are prohibited from charging fees for the 6.5-hour instructional day, and additional funding is granted to providers for assisting low-income students. To this end, providers are required to classify enrolled students into two categories based on their income: a child is eligible for Category One if they or their family participate in the Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), Medicaid, Temporary

Assistance for Needy Families (TANF), or the Childcare and Parent Services (CAPS) program and are classified as Category Two otherwise. Providers are prohibited from charging fees for meals or transportation for Category One students.

Despite the nearly identical provisions for SBPK and non-SBPK providers, a few practical differences exist that may influence parental choice. In addition to SBPK sites requiring applicants to reside in the school district, families may be limited in the number of school-based sites to which they can apply. In the metro-Atlanta area school district we study (henceforth “the District”), parents may only apply for a single school-based site. Meanwhile, there is no limit on the number of non-SBPK sites to which families can apply. In practice, families may apply to both.

The rate at which transportation is provided is another key difference between SBPK and non-SBPK sites. While providers cannot charge fees for transportation to Category One students, offering transportation is optional. According to DECAL’s public data on providers, almost all school-based sites (98.7%) in the District provide transportation to and from school. Meanwhile, only a handful (5.5%) of non-school-based sites in the District do the same, a difference likely arising from the availability of existing busing infrastructure at school-based sites. DECAL compensates providers for transportation at a rate of $16.50 per month per eligible child. This rate may be commensurate for a larger-scale, efficient busing system, but implementing transportation could be economically infeasible for sites where few children would use or need transportation.

The stark difference in the rate at which school-based and non-school-based providers implement transportation raises some concerns about the equity of access to universal pre-K. Transportation bears direct costs in the form of fuel, vehicle maintenance, or public transportation fees. It also presents indirect costs; time spent taking children to school is time that could have been spent working or engaging in some other activity. The fact that some families may have one or no vehicles or no ready access to public transportation exacerbates the problem. Assuming the extent to which these costs are relevant varies based on income, low-income families could effectively have fewer choices even among programs with no out-of-pocket costs.

The number of children seeking entrance to SBPK programs frequently exceeds the number of seats available. DECAL does not dictate how programs allocate these seats in over-subscribed programs, leaving room for variation in enrollment processes. For example, Peisner-Feinberg et al. (2013) surveyed programs across the state during the 2012–13 year and found that, while most (77%) use a first-come-first-served system, the remaining 23%, including the District, use a lottery to determine assignment. In the District, enrollment lotteries for attendance during the next academic year occur each spring. To participate, a child must be four years old by September 1st of the calendar year in which they apply and reside within the District's attendance boundaries.

2.2. Prior Evidence

Past research shows ample benefits from high-quality early childhood education programs. Evidence suggests that interventions early in life are more effective at producing equitable outcomes than those that occur in adulthood (Currie, 2001). Randomized experiments, like the High/Scope Perry Preschool Program in the 1960s and the Carolina Abecedarian Project (CAP) in the 1970s, demonstrated extraordinary value for participating children from low-income families.

Attendees of these programs enjoyed benefits that lasted well beyond their years in school. Being selected to participate in the Perry Program raised children's lifetime earnings by about $200,000 (Belfield et al., 2006). Male CAP participants earned $20,000 more at age 30 and female CAP participants were more likely to be employed at 30 (Garcia et al., 2020). Children who were selected to participate in the Perry Preschool Program spent significantly less time in prison or probation and received about $3,000 less in government assistance. Meanwhile, children who received services from the CAP had greater earnings, were more likely to be employed, and were less likely to be arrested in adulthood (Garcia et al., 2020). Benefits extended beyond the children; parents of CAP participants saw increases in earnings between $7,000 and $14,000. Indeed, the Perry Preschool Program and Carolina Abecedarian Project respectively generated $12.90 and $7.30 of public benefit for every dollar invested (Belfield et al., 2006; Garcia et al., 2020).

Among the earliest of the public interventions in early childhood care is Head Start, which began in the 1960s and sought to provide education and health support to poor children between ages three and five, as well as providing support to their parents. Children who attend Head Start have greater achievement in early elementary school (Deming, 2009) and are more likely to graduate from high school (Ludwig & Miller, 2007). The benefits seem to be greatest for children of below-average initial ability (Deming, 2009). Participants are also more likely to be immunized (Currie & Thomas, 1995) and less likely to die from preventable causes (Ludwig & Miller, 2007).

The large benefits exhibited by the experiments in the second half of the 20th century focusing on families with low incomes, and to a lesser extent Head Start, have generated widespread advocacy for public implementation of early childhood education and care. However, the benefits of universal (no income basis for admission) pre-K programs, like Georgia's, are less clear. Reviewing thirty studies on universal early childhood education conducted between 2005 and 2017, Huizen and Plantenga (2018) find that only one in three studies show positive effects while one in six show negative effects. Even among studies observing the same type of outcome, results can be mixed. For example, Durkin et al. (2022) find evidence that attendees of the Tennessee Voluntary Pre-K Program may later have worse behavior than non-attendees, while studies of other programs show behavioral improvements (Belfield, 2006; Chor et al., 2016). Belfield (2006) even finds that non-attendees benefit from the presence of attendees in a kindergarten classroom.

The common result is that children who participate in any type of pre-K tend to perform better on achievement tests or cognitive measures shortly after the pre-K year (Chor et al., 2016; Currie & Thomas, 1995; Lee et al., 1990; Lipsey et al., 2018). However, these effects are commonly shown to diminish and perhaps disappear completely over time[11]. Creating long-term changes in children's cognitive ability is difficult in the first place (Currie, 2001), and elementary schools might not be taking advantage of the greater preparation of pre-K attendees (Lipsey et al., 2018). Fading quickly, the academic benefits of pre-K can disappear by first or second grade (Lee et al., 1990; Lipsey et al., 2018). One study found that the rate of decay varies based on the characteristics of students. For example, Currie & Thomas (1995) observe Black students seeing the greatest decreases in impact over time—suggesting differences in program delivery or in the types of schools attended by students of different races after early learning.

The uneven findings from research on universal pre-K lies in stark contrast to the preponderance of evidence supporting targeted, high-quality ("model") programs like the Carolina Abecedarian Project and the Perry Preschool Program. Cost-benefit analyses of these programs illustrate the disparity. The benefit of universal pre-K is generally found to be in the

[11] One dissenter is Huizen & Plantenga (2018), whose meta-analysis of universal early childhood education studies suggests no fade out.

range of $2 to $4 for each dollar invested (Bartik et al., 2012; Karoly, 2016). This clearly departs from estimated benefits as high as $17 for model programs (Karoly, 2016). Previous explanations of this discrepancy have noted differences in the funding and intensity of model and public pre-K programs (Currie, 2001; Huizen & Plantenga, 2018). The Carolina Abecedarian Project, for example, spent more than $20,000 per child each year adjusted for inflation (Arnold Ventures, 2017), about four times as much as the Georgia Pre-K Program[12]. It also involved children from eight weeks old to five years old, had no more than six children to a teacher, and operated year-round (Arnold Ventures, 2017). Meanwhile, the Georgia Pre-K program, like other state-funded universal pre-K programs, includes only four-year-olds and permits no more than eleven children per teacher.

Even if programs today had the same funding and intensity, it is possible that their measured benefits would still pale in comparison to those of past programs. The effect estimated depends on the counterfactual—the education a child would have received had they not attended pre-K—and some argue that this comparison is changing. Lipsey et al. (2018) makes this argument, contending that children today have more educational resources, like the internet, at home. Furthermore, Karoly et al. (2016) notes that children in the past were less likely to attend any pre-K program. Students from low-income families, who may have less ability to learn at home, tend to benefit the most from universal pre-K (Huizen & Plantenga, 2018).

Another difference between the early model interventions and universal pre-K that may contribute to the overwhelming positive effect of the former is that the model interventions were targeted to disadvantaged children only. The Perry Preschool Program was targeted to disadvantaged African American students living in Ypsilanti, Michigan (Heckman et al., 2010a,

[12] National Institute for Early Education Research, 2018

2010b) while the Carolina Abecedarian Project targeted disadvantaged and predominantly African American students in Chapel Hill, North Carolina (Garcia et al., 2020). There is no such restriction in universal pre-K programs.

2.3. Data

This study centers around admission lotteries that took place in the District between 2012 and 2018. As stated above, in the District, enrollment lotteries for attendance during the next academic year occur each spring. Site rosters and waitlists help identify winners and non-winners respectively. GA Pre-K sites submit rosters of all enrolled students four times a year to DECAL. Likewise, those sites also send an updated list of students who are actively waiting for spots in the site four times a year. In other words, providers are responsible for maintaining the waitlist by identifying students who no longer wait. While these waitlists do not explicitly identify lottery non-winners, it does record when students enter the waitlist for each site. In the District, full sites accept late applications until August 31; however, these sites only process the applications after exhausting the waitlist. While the ideal strategy would be to identify students who entered the waitlist just after the lottery in spring (since they are the most likely to be lottery non-winners and not late applicants), the earliest date of entry sites can select when adding students to the waitlist is July 1. Therefore, we assume that students lost a lottery if they entered the waitlist on that date.

Students who participate in a lottery and lose may later appear on the roster of another GA Pre-K site. In addition, a student could be removed from the waitlist if a spot opens at their preferred site and causes them to leave the waitlist. Otherwise, they can enroll in another school-based or non-school-based site. In some cases, both happen: a student enters a non-school-based site in Georgia's Pre-K Program, but later enrolls in the site for which they were waitlisted. With that in mind, for questions (1) and (2), We compare students who won the lottery to those who did not win the lottery and never enrolled in any site in Georgia's Pre-K Program. We define lottery sites as sites that had at least one non-winner in a given year. A student is defined as a lottery winner if they appear on a roster for a lottery site but never appear on that site's waitlist.

Since our control group consists of lottery non-winners who don't go to any GA pre-K site, SBPK or non-SBPK, it is important to explore the choices made by them. Table 7 shows the number and percentage of lottery non-winners who attend each type of pre-K site. The most common outcome for children who lose a lottery is to not attend any GA pre-K site, accounting for nearly half of all non-winners (49.59%). For the other half of students who remain in a GA pre-K site, the typical choice (28.38%) is to enroll in a non-SBPK; this constitutes more than half (56.30%) of non-winners who attend GA pre-K sites. Some non-winners (18.94%) do later attend an SBPK, with the majority (11.62%) attending the SBPK for which they originally lost a lottery. Finally, a small number of students attend multiple sites. The most frequent (2.42%) situation in which this occurs is when a child attends both a non-SBPK and their preferred SBPK over the course of a year. One takeaway from Table 7 is that losing an enrollment lottery doesn't necessarily preclude attendance of a pre-K in Georgia's Pre-K Program, both for school-based and non-school-based sites. Considering this, Appendix Table B1 examines the intensive margin of school attendance, comparing students who did not win the lottery with all GA Pre-K attendees, even in non-lottery sites, who were never on a waitlist.

To measure the effects of attending SBPK on K-12 outcomes, we use administrative panel data on students who attended public school in the District. In addition to demographic information like gender, race/ethnicity, English learner status, and free or reduced-price meals eligibility, we also observe key outcomes: absences, disciplinary infractions, and performance on the Measures of Academic Progress (MAP) formative assessments in math and reading. Using these data, we follow students for several years after exiting pre-K and entering the District.

Table 7. Non-Winner Decisions

	# non-winners	% non-winners
No Pre-K Observed	3,285	49.59
Non-SBPK	1,880	28.38
Preferred SBPK	770	11.62
Other SBPK	485	7.32
Non-SBPK & Preferred SBPK	160	2.42
Non-SBPK & Other SBPK	25	0.38
Preferred SBPK & Other SBPK	19	0.29

Note: If a student loses a lottery for an SBPK, that SBPK is considered "preferred" by that student

2.4. Methods

A challenge to estimating the impacts of school-based sites in Georgia's Pre-K Program is that families decide both (i) whether or not to seek admission into a specific program for their child and, (ii) conditional on whether they are offered admission to the desired program, what early learning program (if any) they choose to enroll their child into. Figure 1 illustrates the many choices parents face with respect to their child's early education. If parental choice over programs is influenced by factors that also drive student outcomes (e.g., family income), then a simple comparison of outcomes for students who attend a SBPK program with those who do not attend any Georgia Pre-K Program site would conflate the true impacts of the program with the characteristics of the children and their families.

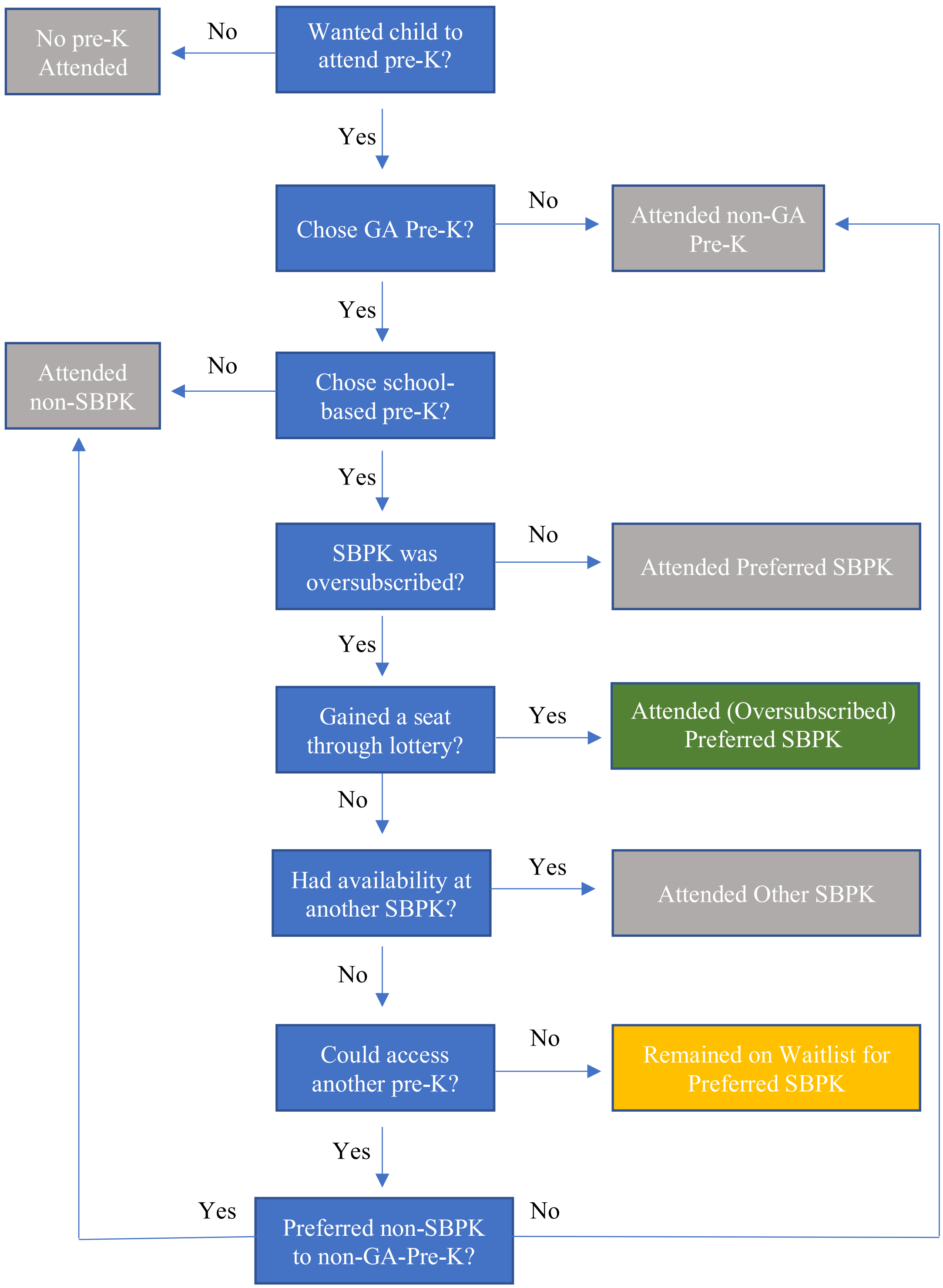
To mitigate potential bias from parental decisions to apply for admission to a SBPK program, we limit our analyses to students whose parents applied for admission to an over-subscribed SBPK program in the District and were thus participants in an enrollment lottery. Given that admission offers are randomly assigned to lottery participants, the characteristics of lottery winners (who are offered admission) and the characteristics of non-winners (who are not initially admitted) should be equal on average and thus eliminate any bias from unobserved family characteristics associated with the decision to apply for a slot in a SBPK program. Figure 7 highlights these groups in green and yellow respectively.

While the characteristics of students should be balanced in winners and non-winners within a lottery, the characteristics of students may not be balanced between lotteries. In other words, while the winners and non-winners in the same lottery may be similar on average, they may be different across different lotteries. To this end, we use lottery fixed-effects which enables comparison of students within lotteries, controlling for systematic differences between students across lotteries.[13] We also use a year fixed-effect for the year in which an outcome was measured. In doing so, we account for potential variation over time in outcomes for all students that is unrelated to attendance of a SBPK.

While including student characteristics in the model would be unnecessary in a fully randomized lottery, our sample is only partially randomized. For it to be fully randomized, among the lottery participants, the decision to go to a SBPK site would need to be unrelated with student characteristics. This is an untenable assumption because the lottery non-winners have a choice to go to a different GA Pre-K program, e.g., a non-SBPK. We control for student demographic characteristics which helps mitigate bias arising from selection into the control group.

[13] The lottery fixed effect is defined as a site-year combination. If a school was observed having a lottery five years in a row, it would generate five different lottery fixed effects.

Figure 7. Example Decision Tree for Parents



Equation 1 shows the model to be estimated. $Y_{ip,g=g^0}$ is the outcome in selected grade $g = g_0$ for student i who participated in lottery p. win_{ip} is the treatment indicator and equals one if student i won enrollment lottery p. X_i is a vector of control variables including race, gender, FRPM status, and ESL status. $\delta_{t_i(g^0)}$ is a fixed-effect for the year t_i in which student i had an outcome for grade g^0, while λ_p is a lottery fixed-effect. The estimated coefficient of interest β_1 is the effect of winning the lottery and attending an oversubscribed SBPK on our outcome of interest.

$$(1): \quad Y_{ip,g=g^0} = \beta_0 + \beta_1 win_{ip} + \beta_2 X_i + \delta_{t_i(g^0)} + \lambda_p + \epsilon_{ipg}$$

To assess heterogeneity of the benefits from pre-K by family income, we estimate a version of equation (1) which includes an interaction of the treatment indicator with an indicator for a student having ever qualified for FRPM. In equation (2), $win_{ip} \times FRPM_i$ is the interaction term between the treatment indicator and FRPM status. The average marginal effect of winning a lottery and attending for FRPM-qualifying students can be calculated by adding the coefficient on the treatment indicator and the interaction term (i.e., $\gamma_1 + \gamma_2$). On the other hand, the coefficient to the interaction (γ_2) is the difference in marginal effect of an FRPM student winning the lottery and attending an SBPK compared to a non-FRPM student winning the lottery and attending the same. A large and significant interaction coefficient would suggest that attending a SBPK is more important for one group than the other. Because a common finding in the pre-K literature is that disadvantaged students tend to benefit more from pre-K attendance, one might expect the interaction to be positive.

$$(2): \quad Y_{ip,g=g^0} = \gamma_0 + \gamma_1 win_{ip} + \gamma_2 win_{ip} \times FRPM_i + \gamma_3 X_i + \delta_{t_i(g^0)} + \lambda_p + \epsilon_{ipg}$$

2.4.1. Limitations

Restricting the analysis to participants in enrollment lotteries does not eliminate potential group differences from subsequent family decisions about where to enroll their child. Students who win a SBPK lottery are eligible to attend but may choose not to. If the attendance decision is correlated with factors that drive student outcomes, it could lead to biased estimates of the impact of SBPK attendance. For example, suppose that more affluent families frequently decide to send their child to a private early-learning program outside of Georgia's Pre-K Program, even when they win the school-based admission lottery, whereas lower-income families cannot afford non-subsidized private alternatives and almost always enroll their child in a SBPK site if they win the lottery. Assuming that more affluent families can also provide additional educational support that raises student outcomes, this would make it look like the SBPK is less effective than it truly is. Similarly, our control group consists of lottery non-winners who do not attend any GA Pre-K. If more affluent families who lose the school-based lottery are more likely to find a non-SBPK site for their child (rather than no formal child care at all), they would be underrepresented in our control group, which could depress outcomes for the control group and overstate the efficacy of attending SBPK.

A second concern is that we do not observe the early childhood educational choices of students that do not attend any site in Georgia's Pre-K Program. While our data covers all public and private sites that are part of the system overseen by DECAL, families have a variety of options (of varying quality) outside of Georgia's Pre-K Program. For example, some early-learning centers in the District that are generally perceived as high-quality, like Montessori schools, are not administered by DECAL. Students who attend these schools could raise the average readiness for the control group. This, in turn, would lower the size of the effect we estimate. On the other hand, children who do not win the SBPK lottery and do not attend any site in Georgia's Pre-K Program could end up in informal childcare settings, such as staying with a neighbor or relative, that may or may not provide strong early-learning opportunities.

Our later analyses attempt to discern the effect of gaining a seat in an oversubscribed SBPK for students who qualify for free or reduced-price lunch, a rough measure of poverty. While this is an important analysis from an equity perspective, it also partially addresses the concern raised previously. Namely, if we assume that higher income families have greater access to other high-quality sites outside Georgia's Pre-K Program than lower income families, children in the latter group would be more likely to have no formal early education. In this case, FRPM-qualifying non-winners would be less likely to attend such a site, and the effect measured among FRPM students should better capture the effect of attending a SBPK versus attending no GA Pre-K.

Third, our analytical strategy relies on comparing winners and non-winners within oversubscribed schools. Our estimates measure the effect of attending an SBPK relative to no attendance of any site in Georgia's Pre-K Program. The extent to which our findings apply to pre-K sites that are not over-subscribed is not clear. The level of oversubscription at pre-K sites is highly likely to be nonrandom. Demand for "good" schools could cause effective pre-K sites to be oversubscribed. Thus, one cannot necessarily extend our findings to school-based sites that are not over-subscribed. In the same vein, our results come from only one school district, and may not be generalizable to other school districts in Georgia or elsewhere.

The fourth issue pertains to the likelihood that a student enrolls in the District in later years and whether winning a lottery affects that likelihood. Our data on elementary school outcomes only covers students who were enrolled in the District, and some students may be more likely to leave than others. This may be a problem if the types of students who are more likely to leave also tend to get a different level of benefit from attending pre-K.

2.5. Results

2.5.1. Effect of SBPK Attendance on Academic Achievement in Elementary School

We begin by estimating the effect on academic achievement in Kindergarten and beyond from winning an enrollment lottery and attending an oversubscribed SBPK program. Academic performance is measured using national percentile ranks in math and reading from the Measures of Academic Progress (MAP) exam. Students in the District take the exam at each grade level during early fall, winter, and late spring. This structure permits evaluating how well prepared a student enters a grade and how their performance evolves over that school year. The MAP exam taken during the fall of kindergarten is of particular interest. Such timing permits little instruction prior to testing, meaning that experiences before kindergarten should drive differences in this score.

Table 8 depicts the estimated effect of attending an oversubscribed SBPK on national reading and math percentiles by grade and test timing following equation (1). In short, it answers the following question: If the average student who lost a lottery (and then never attended any GA Pre-K site) had instead won their lottery and attended, how would we expect their national percentile to change?

Pre-K attendees entering kindergarten score 5.68 and 5.78 percentiles higher on the fall reading and math exams, respectively, than non-attending peers who lost an attendance lottery and did not go to any GA Pre-K. A near six percentile difference is quite large, suggesting that attendees tend to be much better prepared for kindergarten. However, this effect is cut almost in

Table 8. Effects of Pre-K Attendance on MAP National Percentile Scores, by Grade, Subject, and Test Timing

	Reading			Math		
	Fall	Winter	Spring	Fall	Winter	Spring
Kindergarten	5.676***	3.559**	1.962	5.779***	3.346*	2.233
	(1.050)	(1.240)	(1.315)	(1.099)	(1.309)	(1.350)
N	2613	2575	2536	2632	2574	2531
Grade 1	-0.086	0.929	0.074	0.646	-0.479	-0.319
	(1.006)	(1.005)	(1.034)	(1.064)	(1.021)	(1.049)
N	4072	3994	3901	4079	3998	3907
Grade 2	-0.222	-0.427	-1.013	-0.243	-0.453	-0.773
	(0.792)	(0.841)	(0.914)	(0.885)	(0.939)	(1.076)
N	5565	5445	4641	5584	5439	4638
Grade 3	-0.739	-0.390	-1.285	-1.266	-1.566*	-0.855
	(0.885)	(0.854)	(1.010)	(0.818)	(0.797)	(0.921)
N	5890	5814	4418	5916	5811	4412
Grade 4	-1.181	-1.861*	-0.789	-2.018*	-1.279	-2.631*
	(0.925)	(0.930)	(1.148)	(0.866)	(0.839)	(1.065)
N	4877	4790	3395	4903	4795	3396

Note: *, **, and *** represent significance at the .05, .01, and .001 levels.

half after a semester of instruction in kindergarten: SBPK attendees score just 3.35 percentiles higher in math and 3.56 percentiles higher in reading on the winter test than non-winners who did not attend any GA Pre-K. By the test at the end of the spring, point estimates have been cut nearly in half once more, and are marginally insignificant (at the 5% level). The downward trend of the effects which began in kindergarten continue through first grade, where negative, but insignificant, point estimates emerge. By second grade, all point estimates are negative, a situation which never reverses in further grades. For the 3rd grade winter math test, the 4th grade winter reading test, and the 4th grade fall and spring math tests, estimates are negative and significant, which may suggest detrimental effects from attendance of a school-based pre-K.

The gradual decrease in the positive effect of SBPK can also be observed in Figures 8 and 9 for math and reading respectively. The height of the bar represents the expected difference in math percentile rank between students who win an enrollment lottery and attend a SBPK site and students who do not win a lottery at the same site and end up at a non-GA Pre-K early learning center or in no formal pre-K. Shaded bars represent estimated differences in outcomes that we can be 95% confident are not zero.

At first glance, the emergence of statistically significant negative impacts of SBPK attendance on test scores in 4th grade is surprising. However, significant negative effects from attending universal pre-K are not unheard of. Durkin et al. (2022) find some negative effects in later grades when evaluating Tennessee's Voluntary Pre-K Program, and Huizen and Plantenga (2018) indicate that one in six evaluations of universal pre-K programs show significant negative effects. However, our results may also suffer from the sources of bias discussed in the limitations section. In particular, some students who lost a lottery and never attended a site in the GA Pre-K program could go to a high-quality, non-GA-Pre-K private program instead. Because the data covers only GA Pre-K sites, such students appear to have never attended pre-K and therefore enter the control group. Likewise, students who attend high-quality non-GA-Pre-K options may perform better academically regardless of pre-K. If the effect of attending pre-K fades for both groups, a difference in later grades could reflect only the differences in group characteristics. While this issue may be affecting the level of our estimates, it is unlikely to be changing their pattern. Overall, it seems that attendance of an oversubscribed SBPK confers a significant boost to students when they enter kindergarten that fades rapidly as non-winning peers catch up.

Figure 8. Effect of School-Based Pre-K Attendance on MAP Percentile in Math

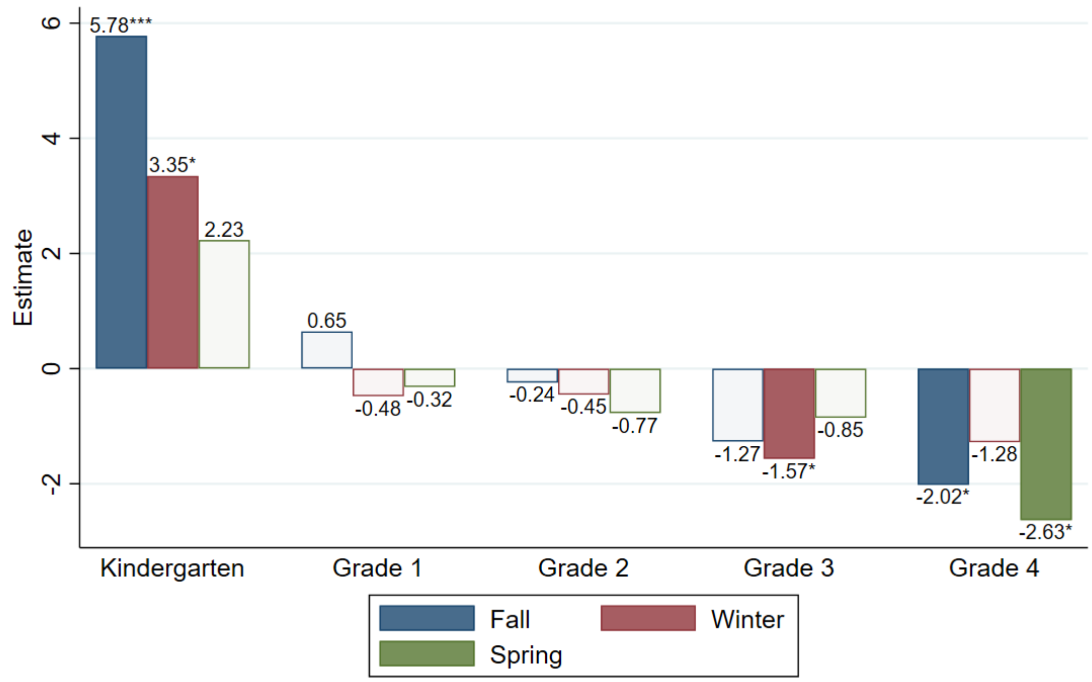
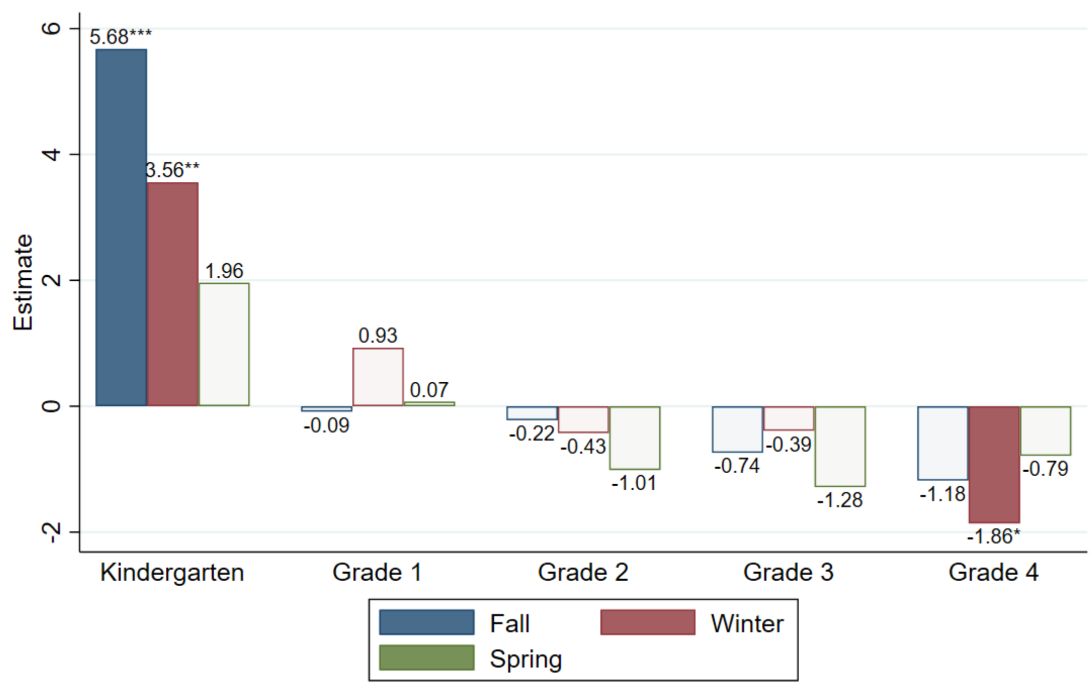


Figure 9. Effect of School-Based Pre-K Attendance on MAP Percentile in Reading



2.5.2. Differential Effects of Pre-K Attendance for Low-Income Students

Evidence suggests that early childhood education can play a significant role in the development of children from disadvantaged backgrounds, namely those from low-income families (Currie, 2001). Universal pre-K is in part organized around the belief that an early intervention can have large effects for that group by reducing the disparity in resources available to children from different economic backgrounds. To better understand the role of early childhood education for low-income students, we repeat the previous estimation while including an interaction of the treatment indicator with FRPM eligibility following equation (2). Table 9 presents the results of that estimation.

The coefficient for win is the marginal effect of SBPK attendance for non-FRPM qualifying students. A positive significant effect for this group only emerges for the fall reading test in kindergarten, showing a 3.72 percentile higher score for SBPK students compared to students who applied but were not granted admission. Point estimates are positive, but insignificant, for four out of the other five kindergarten tests. However, in grade 1 onwards, we often see a statistically significant negative effect of SBPK on non-FRPM students.

The coefficient on $win\ x\ FRPM$ is the difference in the marginal effect of winning between FRPM and non-FRPM students. Coefficients to this interaction term are always positive and often significant. This indicates that the marginal effect of SBPK attendance on test percentiles for students who ever qualified for FRPM is measurably greater than those students who never qualified for FRPM. Taking the results from Tables 8 and 9 together, the large, positive aggregate effects shown in Table 2 seem to be driven by FRPM students.

Two explanations are plausible for the pattern of results exhibited by both subject tests. Recall that, because of the limitations of our data, we are unable to distinguish between going to

Table 9. Effects of Pre-K Attendance on MAP National Percentile Scores, by Grade, Subject, and Test Timing

		Reading			Math		
		Fall	Winter	Spring	Fall	Winter	Spring
Kindergarten	*Win*	3.722*	1.016	-0.833	3.542	1.065	0.953
		(1.784)	(2.187)	(2.176)	(1.957)	(2.200)	(2.024)
	Win x FRPM	3.071	3.984	4.383	3.513	3.578	2.006
		(2.139)	(2.609)	(2.681)	(2.297)	(2.661)	(2.618)
N		2613	2575	2536	2632	2574	2531
Grade 1	*Win*	-3.330*	-2.493	-3.360*	-1.800	-3.814*	-2.562
		(1.584)	(1.610)	(1.577)	(1.693)	(1.567)	(1.596)
	Win x FRPM	5.151**	5.429**	5.461**	3.884	5.282**	3.559
		(1.983)	(2.007)	(2.040)	(2.118)	(1.987)	(2.055)
N		4072	3994	3901	4079	3998	3907
Grade 2	*Win*	-2.103	-2.804*	-1.546	-2.197	-1.764	-2.348
		(1.267)	(1.323)	(1.435)	(1.302)	(1.316)	(1.415)
	Win x FRPM	2.765	3.474*	0.767	2.869	1.917	2.264
		(1.593)	(1.674)	(1.803)	(1.705)	(1.760)	(1.964)
N		5565	5445	4641	5584	5439	4638
Grade 3	*Win*	-2.318	-2.029	-2.122	-2.775*	-3.344*	-2.668
		(1.425)	(1.372)	(1.549)	(1.321)	(1.341)	(1.465)
	Win x FRPM	2.149	2.611	1.776	1.873	1.792	2.116
		(1.710)	(1.664)	(1.886)	(1.580)	(1.590)	(1.763)
N		5890	5814	4418	5916	5811	4412
Grade 4	*Win*	-4.048**	-4.902***	-3.506*	-4.735***	-4.337***	-5.101***
		(1.287)	(1.330)	(1.591)	(1.261)	(1.282)	(1.482)
	Win x FRPM	3.959*	3.819*	2.570	3.460*	3.944*	3.113
		(1.614)	(1.667)	(1.970)	(1.553)	(1.556)	(1.831)
N		4877	4790	3395	4903	4795	3396

Note: *, **, and *** represent significance at the .05, .01, and .001 levels.

a non-GA-Pre-K site (like many Montessori schools) and not going to any pre-K site. Our control group, then, contains both children who don't attend any pre-K and those who attend a non-GA-Pre-K program. We can expect that non-FRPM students are more likely to be able to afford non-GA-Pre-K options and hence are less often classified correctly as not having attended any pre-K, implying that the "true" effect is being captured less frequently among non-FRPM students. Second, early interventions for students from low-income backgrounds could benefit those students beyond direct education. Entering education at age four rather than age five may help remedy resource disparities between high and low-income children, for instance by providing nutritious meals or by giving parents, especially mothers, greater flexibility in employment.

2.5.3. Effects of SBPK Attendance on Absences and Disciplinary Infractions in Elementary School

In the previous section, we showed that attending an oversubscribed SBPK yields large gains in math and reading percentiles at the start of kindergarten which decay as students entered later grades. Prior research has shown that high-quality pre-K programs can yield benefits beyond just helping students score higher on tests, however. School-based Pre-K in the District seeks to promote social-emotional well-being for students in addition to enhancing their educational achievement. We don't have any direct measures for social-emotional well-being. However, given prior literature's findings about non-test score effects, we broaden our analysis to examine two other measures: later attendance and disciplinary conduct. We generate estimates once again by comparing the outcomes of winners and non-winners within lotteries.

Table 10 shows the effect of oversubscribed SBPK attendance on attendance and the number of disciplinary infractions in each grade. We estimate equation (2) to separately identify

Table 10. Effects of Pre-K Attendance on Discipline and Attendance, by Grade

	Kindergarten		Grade 1	
	# Disciplinary Infractions	Days Absent	# Disciplinary Infractions	Days Absent
Win	-0.003	-0.256	0.002	-0.485*
	(0.005)	(0.224)	(0.004)	(0.214)
Win x FRPM	-0.004	-0.150	-0.018	-0.236
	(0.010)	(0.289)	(0.013)	(0.302)
Constant	0.044***	4.853***	0.022	4.404***
N	9802	9675	9163	7957

	Grade 2		Grade 3	
	# Disciplinary Infractions	Days Absent	# Disciplinary Infractions	Days Absent
Win	-0.002	-0.300	-0.011	-0.737*
	(0.004)	(0.249)	(0.006)	(0.309)
Win x FRPM	0.021*	-0.444	0.031	0.347
	(0.009)	(0.360)	(0.020)	(0.436)
Constant	0.035**	4.096***	0.050***	4.802***
N	8828	5835	7774	3912

	Grade 4	
	# Disciplinary Infractions	Days Absent
Win	0.003	-1.022**
	(0.003)	(0.391)
Win x FRPM	0.002	0.282
	(0.009)	(0.659)
Constant	0.021**	4.416***
N	5928	1974

Note: *, **, and *** represent significance at the .05, .01, and .001 levels.

the effect on FRPM and non-FRPM students. When it comes to disciplinary infractions, we find a positive interaction term for discipline in grade 2 implying a positive marginal effect of SBPK attendance on disciplinary infractions for FRPM-qualifying students. This, however, is the lone significant result for discipline. In general, we do not find a relationship between attending a SBPK and disciplinary infractions. Because a student's number of disciplinary infractions is only a loose measure of their overall social and emotional competency, these results are not necessarily indicative of the ineffectiveness of SBPK in the District for nurturing social-emotional learning.

On the other hand, winning a lottery for and attending an SBPK does appear to significantly decrease the number of days for which a student is marked absent in first grade, third grade, and fourth grade. Point estimates are negative but insignificant for kindergarten and second grade. FRPM status does not appear to moderate this effect. Taken as a whole, the relationship here is modest, with the average winner attending roughly three more days of school between kindergarten and fourth grade.

2.5.4. Student Characteristics and Pre-K Enrollment Behavior for Lottery Non-Winners

Certain characteristics of students are predictive of whether and where a student goes to pre-K. We use a multinomial logit model to estimate the marginal effect of membership in various subgroups on the relative likelihoods for different types of pre-K attendance among children who enter lotteries for over-subscribed SBPK sites but do not win the lottery and thus are not offered admission. Table 11 provides the coefficients from that model, which should be interpreted as the marginal effect of the given subgroup on the log-odds of attending SBPK, Non-SBPK, or both, relative to not attending any pre-K. Further discussion below interprets the coefficients as odds rather than log-odds, which one can obtain by exponentiating the coefficient. These exponentiated log-odds are also listed in the table as the odds-ratio.

Table 11. Marginal Effect of Subgroup Membership of Nonwinners on Odds of Pre-K Attendance Type

	SBPK	Non-SBPK	Both
	Odds ratio [Log odds] (Standard error)	Odds ratio [Log odds] (Standard error)	Odds ratio [Log odds] (Standard error)
Female	1.162* [0.150*] (0.07)	1.113 [0.107] (0.06)	1.195 [0.178] (0.15)
Black	1.937** [0.661**] (0.20)	2.140*** [0.761***] (0.21)	2.160 [0.770] (0.62)
White	0.730 [-0.315] (0.19)	0.394*** [-0.932***] (0.20)	0.254* [-1.371*] (0.60)
Asian	0.982 [-0.018] (0.21)	1.234 [0.210] (0.21)	1.078 [0.075] (0.62)
ELL	0.537*** [-0.621***] (0.11)	0.838 [-0.177] (0.09)	0.319*** [-1.142***] (0.31)
FRPM	1.487*** [0.397***] (0.09)	1.602*** [0.471***] (0.08)	2.237*** [0.805***] (0.19)
Constant	0.386*** [-0.951***] (0.21)	0.522** [-0.651**] (0.22)	0.055*** [-2.906***] (0.64)
N = 6208			

Note: These are multinomial logit model estimates of the marginal effect of membership in various subgroups on the relative likelihoods for different types of pre-K attendance. The first number for each student characteristic is the odds of attending the program indicated in the column relative to not attending pre-K, the second number is the log odds and the third number is the standard error of the log odds calculated in the multinomial logit model. Subtracting one from the odds ratio allows interpretation as a percentage more or less likely.

English language learners who lose an enrollment lottery are 46.3% less likely to attend a SBPK and 16.2% less likely to attend a non-SBPK relative to not attending GA Pre-K at all. In contrast, FRPM-qualifying students who lose an enrollment lottery are 48.7% more likely to attend a SBPK and 60.2% more likely to attend a non-SBPK than not attending GA Pre-K. White non-winning students are slightly less likely to attend a SBPK and significantly less likely to attend a non-SBPK than no site in Georgia's Pre-K Program, whereas Black non-winner-students are almost twice as likely to attend either a SBPK or non-SBPK than not attending Georgia's Pre-K Program.

Because our data cannot distinguish students who attend a pre-K unaffiliated with the GA Pre-K Program from those who truly do not attend any pre-K at all, it is difficult to interpret these results. White children being less likely to be observed in any pre-K might reflect the use of options outside Georgia's Pre-K Program. In contrast, the finding that English learners who lose a SBPK lottery are more likely to not attend GA Pre-K, rather than attend a SBPK or non-SBPK, may be explained by limited access to ELL services in non-SBPKs and difficulty in obtaining transportation for their children, which could result in staying at home or participating in informal pre-K settings. The choices of FRPM-qualifying students are more difficult to rationalize. Given that few non-SBPK programs offer transportation, it is surprising that FRPM non-winners are relatively more likely to attend a non-SBPK than not attending GA Pre-K at all. The reader should note that these explanations are merely conjecture, as this study does not have data on options outside of Georgia's Pre-K Program. Further research on pre-K in Georgia would greatly benefit from data with more detail on the choices of students who do not attend any pre-K affiliated with Georgia's Pre-K Program, but gathering quality data from a variety of independent early childhood education centers presents a significant challenge.

2.6. Conclusion & Policy Implications

In this paper, we estimated the effects of attending an oversubscribed school-based Georgia's Pre-K Program on achievement, attendance, and discipline in elementary school. Using the results of lotteries for oversubscribed school-based pre-K sites in a metro-Atlanta school district, we compared students who gained a seat through an enrollment lottery and attended a school-based site in Georgia's Pre-K Program to students who did not gain a seat through a lottery and did not go to any site in Georgia's Pre-K Program.

We find that lottery winners enter kindergarten significantly more prepared, around six percentiles, than their non-winning peers as measured by national percentile rankings on the Measures of Academic Progress (MAP) math and reading tests. However, these gains fade by the end of kindergarten, and some negative effects on achievement emerge by grade 4. The negative effects in later grades may be driven by students in the control group who attend options outside of Georgia's Pre-K Program. Measured effects are always statistically significantly greater in grades 1, 2, and 4 for students who qualify for free or reduced-price meals (FRPM) as compared to their non-qualifying peers, suggesting that attending pre-K may be more beneficial for disadvantaged students, a common finding in the early education literature (Lee et al., 1990; Currie, 2001). While winners were no less likely than non-winners to commit a disciplinary infraction in any grade, they did miss about one fewer day of instruction in each grade after Kindergarten.

Importantly, we find that students who qualify for free or reduced-price meals almost always benefit more from winning a lottery for a school-based pre-K and attending. Disadvantaged students who are not in a formal setting may have more limited access to educational resources than their peers, a disparity that pre-K attendance alleviates. Another factor that may be relevant for low-income families is the difference in transportation provision between school-based and non-school-based sites in Georgia's Pre-K Program. While almost all school-based sites offer transportation (which is free for low-income students), almost no non-school-based sites do, and the effects of losing a lottery could be more acutely realized for low-income families who have limited transportation availability.

The limitations of our analysis make us cautious in providing policy recommendations. However, due to the disparities in transportation access across sites, offering transportation-limited students priority at sites which offer transportation could be impactful. In the long-term, additional funding could help non-school-based sites overcome the cost of providing transportation, as they don't have the economies of scale like elementary schools do. Finally, providing additional information to parents could be a relatively inexpensive and potentially impactful way to increase the number of students served. In particular, informing non-winning parents of next steps and other options within Georgia's Pre-K Program reduces the chance that their child does not attend any formal pre-K. Our results give suggestive evidence that this type of intervention could be particularly beneficial if aimed at families with limited language proficiency, as they have a greater barrier to accessing information.

It is possible that providers in Georgia's Pre-K Program are preparing students in ways that we are not measuring. For instance, pre-K may develop its attendees' social-emotional skills. Our null results for impacts on discipline do not support this notion, but they do not necessarily rule it out. Little variation exists in the number of infractions per student, meaning that our model might not be well-suited to detecting a relationship. On the other hand, we do find a consistent, positive relationship of pre-K with later elementary school attendance. This is encouraging insofar as it indicates that attending a school-based pre-K can have a persistent effect on a student, but it is unclear what mechanism drives this decrease in absenteeism. It could also be possible that students who attend pre-K generate positive effects for non-attendees in their classrooms, as Belfield (2006) suggests. For instance, pre-K attendees may be more prepared, or easier to teach, allowing teachers to perform their job more effectively. In theory, these spillovers would raise the readiness of the control group and diminish the estimated effect of attending a school-based pre-K on later outcomes. We cannot conclusively explain the mechanisms driving the patterns shown in this paper.

The broad patterns we find are consistent with previous studies of the efficacy of universal pre-K programs elsewhere: attending a school-based pre-K does prepare students well academically for kindergarten, but these measured benefits do not appear to persist for long. It is not clear why this is the case. One study has suggested that elementary schools might fail to capitalize on the greater academic preparedness of pre-K attendees (Currie & Thomas, 1995). More research is needed to understand the pathways that connect early educational outcomes to those later in childhood.

CHAPTER III. Monopsony in the Market for Remote Work

3.1. Introduction

In recent years, the perfectly competitive model of the labor market has begun to fall from favor relative to models that depict employers exercising significant market power. Monopsonistic labor markets suppress wages for workers to below the competitive wage (Boal & Ransom, 1997), and systemic disparities in monopsony power between less and more populated areas drive a portion of the urban wage premium (Luccioletti, 2023). These markets can be a viable target for policies seeking to remedy the imbalance in bargaining power between employers and employees, and, as a result, much work has gone into measuring labor market concentration and monopsony power. At the same time, the market for fully remote labor has expanded over time, especially during the COVID-19 pandemic, and this trend is not expected to reverse (Chen, Cortes, Kosar, Pan, & Zafar, 2023).

Despite this, no work of which I am aware directly estimates monopsony power in a market for full-time remote work. The closest is Dube, Jacobs, Naidu, and Suri (2020), finding evidence of monopsony power in Amazon's Mechanical Turk (MTurk), an online labor market in which workers can choose to complete simple, on-demand tasks for small payments. As remote work arrangements continue to grow in popularity, understanding the level of monopsony power present in markets for remote labor becomes increasingly important for understanding the generation of wage inequality and creating policies to mitigate it. To remedy this deficiency, I mimic Dube et al. (2020)'s application of Chernozhukov et. al's (2018) double machine learning method to estimate monopsony power in a new setting.

Using a large corpus of job postings for full-time remote labor scraped from online job boards in the United States between 2012 and 2022, I estimate elasticities of posting durations with respect to posted salaries as a measure of monopsony power. I find evidence of monopsony power even in markets for fully remote work, despite ostensibly lower search frictions and concentration than corresponding in-person markets. Likewise, I find that employers of remote labor in high-population commuting zones face more elastic labor supply curves, echoing the results of papers finding lower monopsony power in more populous markets (Azar, Marinescu, & Steinbaum, 2020; Luccioletti, 2023).

3.2. Background

Compared to a *monopoly*, a market with one seller, a *monopsony* is a market with one buyer. First coined by Joan Robinson (1969), the term has come to be used almost exclusively in the context of labor markets in which one or few employers purchase the labor of workers. In the competitive model of the labor market, firms are price takers who have no wage setting power. However, firms in monopsonistic markets enjoy the ability to set wages lower than the marginal product of labor and competitive wage (Boal and Ransom, 1997). Most labor markets around the country are not truly competitive, allowing employers to suppress wages in this way (Naidu, Posner, and Weyl, 2018; Azar, Marinescu, Steinbaum, & Taska, 2020). Luccioletti (2023) estimates up to a 15% reduction in wages from highly monopsonistic markets compared to competitive ones.

If the labor market were perfectly competitive, firms would have no ability to set wages, and an employer lowering the wages of their employees by any amount would result in every employee seeking work elsewhere. In reality, firms do have wage setting power – a 1% reduction in wages does not generally result in a 100% reduction in employment. Put differently, the labor supply curve in the competitive case is perfectly elastic, while labor markets in the real world tend to have various degrees of inelasticity in labor supply. A plethora of reasons could cause real world labor markets to diverge from the canonical competitive one. A labor market can be concentrated, with only a handful of firms hiring a particular type of worker, increasing the bargaining power of firms relative to workers. Workers also face frictions as they navigate the market, with Manning (2003) listing "ignorance, heterogeneous preferences, and mobility costs" as the three most dominant. Workers have imperfect information about what offerings are available or their expected match quality, making it difficult to compare opportunities. Further, each individual may have different preferences for working conditions, job tasks, and non-wage amenities. Finally, moving between jobs is costly, especially if it involves geographic relocation. All told, it is time consuming and costly to find an employer who is a good match for a given worker, and hence people tend to list employment-related happenings like getting or losing a job as some of the most significant events in their lives (Manning, 2011).

The level of monopsony power in a market is an important characteristic for understanding inequality generation and in choosing optimal policies. For instance, the existence of monopsony power explains how a minimum wage increase can increase both wages *and* employment – as Card and Krueger (1994) found – if the wage floor is set between the competitive wage and the monopsonistic wage (Boal & Ransom, 1997). It also explains how equal pay legislation can raise the wages of female workers without lowering their employment (Manning, 2003). Furthermore, more populous labor markets tend to have lower market concentration (Azar et al., 2019) and lower monopsony power (Azar, Marinescu, & Steinbaum, 2020; Luccioletti, 2023), driving a substantial portion of the difference in wages and employment between small and large cities (Luccioletti, 2023).

Simultaneous to the growing interest in monopsonistic markets, remote work arrangements have gained popularity across the United States and are expected to persist (Chen et. al, 2023). Remote work is an amenity that workers value (Haoran, Neumark, & Weng, 2021), and firms offering it enjoy benefits including compensated lower wages (Barrero, Bloom, Davis, Meyer, & Mihaylov, 2022), sometimes higher productivity (Bloom, Liang, Roberts, & Ying, 2015), and lower turnover (Barrero, Bloom, & Davis, 2021).

The market for fully-remote work is structurally different than a traditional in-person labor market. As markets for remote labor have an effectively infinite geographic reach, one might expect it to be less concentrated than a traditional market. Likewise, switching between two fully remote positions based out of two distant labor markets is less costly than switching between two in-person positions in those same markets. These differences evoke a multitude of questions related to monopsony power. For instance, could a competitive market for remote labor provide a means to reduce the wedge in wages between less and more populated area? On the other hand, if that market is competitive, would minimum wage increases reduce employment in remote labor?

While I am unaware of studies measuring market power in labor markets for fully-remote work, a strand of the literature has evaluated Amazon's Mechanical Turk (MTurk), an online labor market where users can perform simple, on-demand tasks in exchange for small payments. Perhaps more than markets for fully remote work, MTurk has attractive characteristics of traditional competitive labor markets like easy entry and exit, no geographic constraints, and high information. Despite this, Kingsley, Gray, and Suri (2015) finds considerable employer concentration in MTurk. Dube et al. (2020) continues further in this vein, using a double machine learning method (Chernozhukov et. al, 2018) to estimate the elasticity of MTurk task listing durations with respect to the task's reward as a measure of the elasticity of labor supply. The duration elasticities they estimate are small, representing an inelastic labor supply, providing evidence of monopsony power.

3.3. Empirical Strategy

3.3.1 Basic Monopsony Model

In the simple model of monopsony (Boal & Ransom, 1997; Manning, 2003) a firm produces output using only labor and faces a labor supply curve relating wages w to quantity of labor supplied L, with the inverse labor supply curve given by $w(L)$. The profit function of the firm is therefore stated in Equation 1 with first-order conditions stated in Equation 2.

$$\pi = Y(L) - w(L)L \quad (1)$$

$$MRP_L = w'(L)L + w(L) \quad (2)$$

By rearranging and dividing both sides by $w(L)$, we arrive at Equation 3, which gives us an expression for the elasticity of wages with respect to labor supply (ϵ_w) facing the firm. Finally, we invert this to get a term for the elasticity of labor supply with respect to wages (ϵ_L). The leftmost term is the difference between the marginal revenue product of labor and the wage paid to labor as a proportion of the wage. Sometimes called the 'exploitation factor' E (Boal and Ransom, 1997), it measures the divergence from the competitive equilibrium where the wage equals the marginal revenue product.

$$\frac{MRP_L - w(L)}{w(L)} = w'(L)\frac{L}{w(L)} = \epsilon_w = \frac{1}{\epsilon_L} \quad (3)$$

$$\Rightarrow \frac{MRP_L - w(L)}{w(L)} = E = \frac{1}{\epsilon_L} \qquad (4)$$

In a perfectly competitive market, the labor supply curve is perfectly elastic ($\epsilon_L = \infty$) and hence the exploitation factor is zero. On the other hand, monopsonistic markets have inelastic labor supply (ϵ_L close to zero) and therefore have a high exploitation factor. In other words, firms in monopsonistic markets can pay workers wages lower than the marginal revenue product of labor. This model illustrates that measuring the elasticity of labor supply is sufficient to measure the market power of a firm in the labor market.

3.3.2 Data

To estimate labor supply elasticity in the market for remote work, I use a large set of job postings from between January 2012 and October 2022 purchased from Emsi, Inc[14]. Emsi scraped or purchased job postings from company websites, job boards, and aggregators, afterward parsing the raw textual data to extract a number of features. Where possible, the algorithm retrieves variables including the company name, job title, location, degree requirements, experience requirements, NAICS industry, SOC occupation, a part-time versus full-time employment indicator, and an internship indicator. They also extract indicators for fully-remote, hybrid, or in-person work. Emsi extracts the posted upper and lower bounds of the salary where available, generating a single salary variable as the average of these bounds.

Finally, each job posting has a date when it was created and a date when it expired. These are critical; an important variable to this analysis is the duration of the posting as calculated as the expiration date minus the posted date. Emsi determined a posting to be expired if the scraper revisited the page and the posting was no longer available OR if the posting had been up for

[14] Emsi merged with Burning Glass Technologies in 2021 to become Emsi Burning Glass and later Lightcast.

longer than 60 days (age-based expiration). The exception to age-based expiration is if multiple advertisements are made for one posting, in which case the expiration date will take the later of the two or more individual advertisements' expiration dates, up to a maximum of 121 days. Due to the policy, nearly half of all calculated durations equal 60 days. As a result, I drop all durations equal to exactly 61 days. I also drop all durations shorter than a week as they are not likely to be accurate given a standard job search timeline. Appendix Figure C1 shows the distribution of calculated durations after the restrictions are applied. Even after the restrictions, the distribution of durations contains significant mass just after the 60-day cutoff.

The data has 326,324,743 job postings from January 2012 to October 2022. I begin by restricting to only job postings with fully-remote work indicated, which lowers the sample to 8,968,960 observations[15]. Dropping all postings with durations of 61 days or fewer than 7 days further lowers the sample to 4,381,409 observations. I then drop all job postings for part-time employment (to 4,381,134) and postings without salary data (to 1,042,122). Finally, I drop all job postings with a salary range, as defined by the difference between the posted minimum and maximum salaries, greater than half the extracted salary. For instance, I would discard a job posting with an extracted salary of $100,000 if the listed salary range was greater than $50,000. My justification comes from an acknowledgement that the extracted salary, being a function of the bounds, is inherently noisy, and the extent of the bounds directly measures that noise. After dropping all such observations, 809,081 remain. Appendix Figure C2 shows the distribution of salary ranges as a proportion of the salary midpoint, while Appendix Figure C3 shows the distribution of salaries after the 0.5 cutoff is applied.

To give a better sense of how the final sample of remote jobs compares to other work in

[15] The reader should not infer that only about 9m of 326m job postings are remote, just that the Emsi algorithm was only able to identify those 9m job postings as remote.

the US, Table 12 reports descriptive statistics for (1) the entire corpus of job postings, (2) job postings identified as fully remote, and (3) the final sample of job postings with all sample restrictions discussed above. The mean advertised salary for the full sample is approximately $57,000, which is close to the mean yearly earnings for US workers. Jobs identified as fully remote tend to have higher advertised salaries, are filled faster, need slightly more experience, and tend to require more advanced degrees. The final sample features slightly higher pay than the full sample of remote jobs but is otherwise similar.

Table 12. Descriptive Statistics by Subsample

	(1)	(2)	(3)
Advertised Salary	57125.72	77,733.57	83,245.54
	(39,168.86)	(45,828.38)	(47,970.13)
Posting Duration	64.95	56.07	56.97
	(39.53)	(32.77)	(30.54)
Min. Years Experience	3.25	3.97	3.49
	(2.70)	(2.89)	(2.65)
Min. Education Level			
High School or GED	41.36%	23.46%	29.71%
Associate Degree	9.55%	6.49%	7.77%
Bachelor's Degree	43.27%	64.01%	56.97%
Master's Degree	4.02%	4.37%	3.48%
Professional Degree	1.80%	1.67%	2.07%
Identified as remote?	No	Yes	Yes
Other sample restrictions?	No	No	Yes
N	326,324,743	9,785,242	809,081

Note: standard deviations in parentheses

3.3.3 Estimating Labor Supply Elasticity

My method for estimating monopsony power in the market for remote labor closely follows Dube et al. (2020), who estimated the same quantity for the online labor market MTurk.

Dube's basic linear specification estimates the log of the task posting's duration as a function of the log of the task's payment and a confounding term. I adapt this in Equation 5 below, where $duration_p$ is the calculated duration of job posting p, $salary_p$ is the posting's advertised salary, and ν_p is an unobserved term that is correlated with both $duration_p$ and $salary_p$.

$$\ln(duration_p) = -\eta \ln(salary_p) + \nu_p + \epsilon_p \qquad (5)$$

The quantity to be estimated is η, called the 'recruitment elasticity' by Dube, measuring the elasticity of a job posting's duration with respect to the posted salary. This 'recruitment elasticity' is a valid measure of the labor supply elasticity insofar that workers supplying labor (to a given posting) more readily results in a shorter time to fill a vacant position. Essentially, a vacancy in a competitive market would be expected to be filled more quickly if the salary is high versus if the salary is low. However, estimating Equation 4 without properly handling ν_p, which is correlated with both duration and salary, would result in a biased estimate of η. For example, a vacancy for a highly-skilled position may have a high advertised salary but be time-consuming to fill.

I attempt to mitigate bias from the presence of confounding factors in the two ways used by Dube. First, I attempt to control for ν_p by estimating a fixed effects model with indicators for year-quarter, job title, and company. This should serve to isolate some of the variation in log duration coming from factors unrelated to salary like workplace desirability, company hiring practices, fluctuations over the business cycle, and differential difficulty of filling positions. This model is shown in Equation 6, where τ_p, δ_p, and γ_p are indicators for year-quarter, job title, and company name respectively.

$$\ln(duration_p) = -\eta \ln(salary_p) + \tau_p + \delta_p + \gamma_p + \epsilon_p \quad (6)$$

Second, I implement Chernozhukov's (2018) double machine-learning estimator as an alternative to the fixed effects model. This approach is a recent development over traditional nonparametric methods like Robinson's (1988) double residual estimator. In short, Chernozhukov's approach leverages the high prediction accuracy of machine learning models, which have typically floundered in economics due to poor inference of causal parameters, to orthogonalize the outcome variable and treatment variable with respect to a high-dimensional set of covariates. Formally, I begin by modeling the unobserved confounder ν_p as a function $g(Z)$, where the functional form of g is unknown but Z_p is a high-dimensional vector of covariates that confound the analysis. The vector Z_p are confounders insofar as $\ln(salary_p) = m_0(Z_p) + V_p$, where m_0 is a non-zero function of Z_p and V_p is uncorrelated noise. The starting point for the new model is shown in Equation 7.

$$\ln(duration_p) = -\eta \ln(salary_p) + g(Z_p) + \epsilon_p \quad (7)$$

Naively estimating this model by using nonparametric or machine learning algorithms to model $g(.)$ directly will result in a biased estimate of η. However, residualizing with respect to $g(Z)$ first, in the style of Frisch-Waugh-Lovell, allows consistent estimation of η. Estimation of this model proceeds by first estimating the conditional expectation functions $d_0(Z) = E[\ln\,(duration_p)|Z]$ and $s_0(Z) = E[\ln\,(salary_p)|Z]$. Then, Equation 8 as follows is estimated for η, where $\widehat{W_p} = \ln\,(duration_p) - \widehat{d_0}(Z)$ and $\widehat{V_p} = \ln\,(salary_p) - \widehat{s_0}(Z)$. The double machine learning estimator is doubly robust and will yield consistent estimates of η if either

$\widehat{d_0}(Z)$ or $\widehat{s_0}(Z)$ are consistent.

$$\widehat{W_p} = -\eta \widehat{V_p} + \omega_p \quad (8)$$

Empirically, the double machine learning estimator requires that these conditional expectation functions be constructed using a machine learning algorithm with high prediction accuracy (R^2). For this case, I use the random forest estimator, an ensemble estimator which is composed of some number of regression trees. These trees form predictions by recursively partitioning the sample space along divisions of values of the covariates Z_p, returning the mean value of the predicted variable at the smallest subdivision. The random forest estimator averages the predicted value from each of the trees in the forest to compute a single final predicted value.

I use a random forest estimator with 100 trees to estimate conditional expectation functions of log salary and posting duration. The covariates Z_p contain several groups of variables. The first are basic extracted information like minimum educational and experience requirements, where available. The second group is a vector embedding with length of 200 generated from training a Doc2Vec model on the cleaned body text of the job postings. Semantically similar postings correspond with closer vector representations, and hence embeddings generated with Doc2Vec help capture information in job postings not extracted by Emsi. The third group are indicators for certain skills being present in the posting. In 2019, Emsi created a dictionary defining 32,000 skills and an algorithm to extract them from unstructured text, generating a list of skills included in each job posting. I include indicators for the presence of each of the 1,118 most common skills, representing the top 5% most common. Finally, the fourth group is a set of indicators for the NAICS 2-digit major industry of the hiring company.

Sample splitting ensures consistency of the double machine learning estimator by removing bias from overfitting. I begin by splitting the data into two[16] equally sized partitions P_1 and P_2. Within P_1 and P_2, I allocate 80% of the observations for training the random forest estimator and 20% for testing. After training the random forest to predict salary and duration using partition P_1, I use it to predict salary and durations in partition P_2 instead. Likewise, a model trained on data in P_2 is used to predict salary and duration in P_1. Table 13 reports the R^2 of the random forest estimators trained on P_1 and P_2 for both variables to be residualized. Despite the imperfect data quality described previously for both durations and salaries, the random forest estimator succeeds in explaining most of the variance in the two variables. After residualizing salary and duration, as the last step of the sample-splitting procedure Equation 8 is estimated for both P_1 and P_2 to yield η_1 and η_2 which are averaged to yield the final coefficient estimate $\hat{\eta}$.

Table 13. R^2 of Random Forest Estimator on Data Partitions

	Partition 1	Partition 2
Duration	0.696	0.694
Salary	0.744	0.746
Training N	334483	334483

Note: Duration and salary predicting using set of covariates including Doc2Vec vector embeddings, indicators for presence of common skills, and indicators for NAICS industry.

3.4. Results

Table 14 shows the results of estimating the duration elasticities using the methods discussed above. The baseline OLS model in Column 1 estimated an elasticity of -0.1064, indicating that a 10% increase in salary would be associated with an approximately 1% decrease

[16] Note: this procedure can be done using an arbitrary number of partitions, but two is common.

in the posting duration. Next, Column 2 estimates the model using OLS once more, but restricts the sample to observations included in the fixed effects model for ease of comparison. Column 3 then includes fixed effects for company name, job title, and year-quarter. The company fixed effect should help remove variation in salary and duration idiosyncratic to the firms; for example, some firms may have, for all positions, relatively high salaries or longer vacancy durations. Meanwhile, the job title and year-quarter fixed effects isolate variation related to the market for a given occupation. Most of the measured relationship from Columns 1 and 2 is erased by adding these fixed effects. Next, Columns 4, 5, and 6 display results from the double machine learning estimator. The baseline double machine learning estimate is -0.0732 but rises to -0.0952 when the sample is restricted to the fixed effects sample. Like in the case of OLS, adding fixed effects attenuates the coefficient; however, the drop is much less severe.

Table 14. Estimation of Elasticities of Posting Duration with Respect to Advertised Salaries

	OLS			Double Machine Learning		
	(1)	(2)	(3)	(4)	(5)	(6)
$ln(salary_p)$	-0.1064***	-0.1674***	-0.0183***	-0.0732***	-0.0952***	-0.0629***
	(0.0015)	(0.0021)	(0.0046)	(0.0016)	(0.0021)	(0.0021)
N	836,208	475,560	475,560	836,208	452,330	452,330
Fixed effects?	No	No	Yes	No	No	Yes

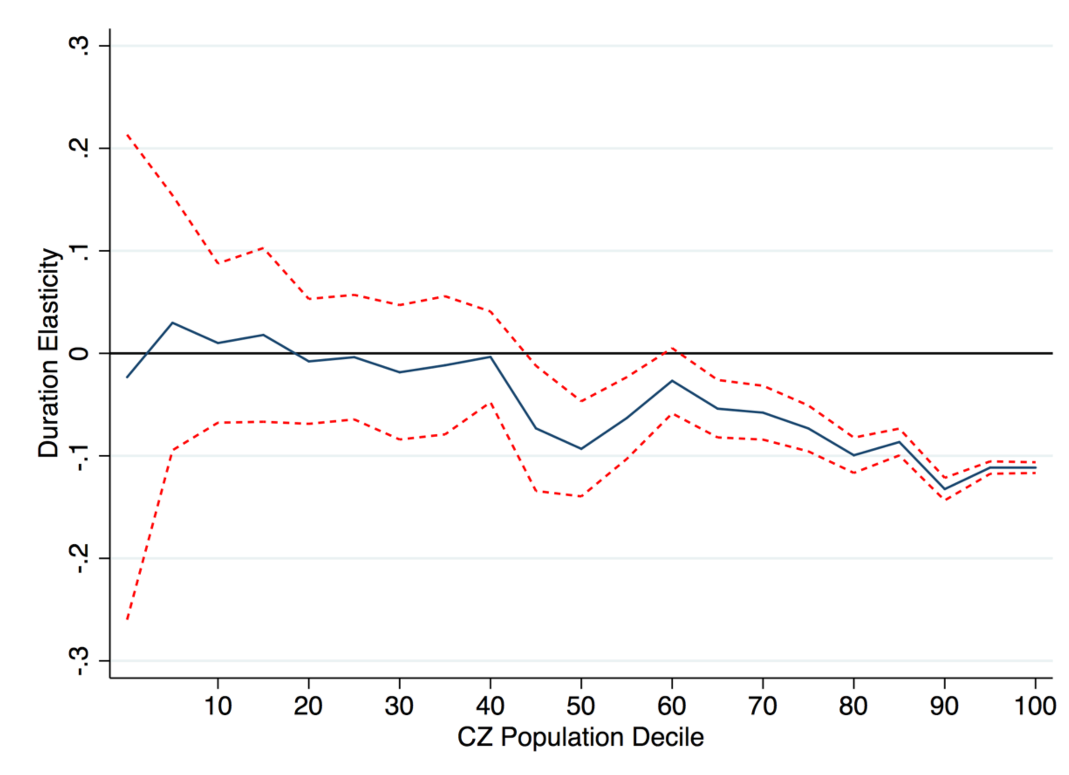
Notes: Fixed effects include company name, job title, and year-quarter.

In each of the specifications, the elasticities I estimate are small. These duration elasticities are remarkably similar to the duration elasticities estimated by Dube et al. (2020), despite that paper analyzing the online labor market MTurk and not a traditional labor market. There, the calculated duration elasticities from the double machine learning estimator were -0.0958 (no fixed effects) and -0.0787 (fixed effects). If the market for remote labor were truly competitive, I would expect to recover much larger estimates of the elasticities. Therefore, I conclude from these estimates that firms hiring for remote labor do exercise some level of monopsony power.

As mentioned previously, a growing body of evidence links high monopsony power in the labor markets of less populated areas to between-city wage inequality (Azar et al., 2019a, Azar et al. 2019b; Luccioletti, 2023). Because fully remote work has, in theory, no geographic constraints, one might expect it to mitigate this source of wage inequality. For this to be true, the population of the nominal area listed on the remote job posting, presumably the area containing the company HQ or branch that is hiring, would need to be unrelated to that area's estimated duration elasticity for remote jobs. In other words, in the market for remote work, each commuting zone should have similar levels of monopsony power.

I test this by estimating a modified version of Equation 8 including interactions of the residualized salaries with ventiles for the populations of associated commuting zones. Figure 10 plots the marginal effect of the residualized salary - the duration elasticity – as a function of these ventiles. Below the 45^{th} percentile, the 95% confidence interval indicates that measured elasticities are statistically indistinguishable from zero. As population increases, the estimated elasticities seem to become more negative. As stated earlier, if fully remote work is truly agnostic to geography, the nominal location of the posting should be irrelevant. If the location of the posting is irrelevant, then likewise there should not be heterogeneity of competition by market size. However, Figure 10 demonstrates that the heterogeneity across markets normally seen in fully in-person work is present for fully remote work as well. While it is unclear what the source of this heterogeneity is, the fact that the same pattern of heterogeneity exists challenges our understanding of the mechanics of remote work.

Figure 10. Elasticities Across the Distribution of Commuting Zone Population



3.5. Discussion

As remote work has grown in popularity across the US, so too have models of the labor market which allow for imperfect competition. Previously, most labor markets around the country have been shown to have some level of monopsony power (Naidu et al. 2018; Azar et al., 2020). Despite this, I am unaware of evidence estimating monopsony power in markets for fully-remote work. One might expect these markets to be competitive due to, for instance, decreased labor market concentration because of the broader geographic scope, or lower search frictions due to no requirement of geographic mobility. Estimating the elasticity of labor supply in MTurk, an online market for crowdsourced work, Dube et al. (2020) showed that even an online market with traditionally attractive qualities can be monopsonistic.

In this paper, I use a large set of job postings from between 2012 and 2022 to answer questions about monopsony power in the market for fully remote labor in the United States. First, as a measure of the elasticity of labor supply and hence the level of competition, I follow Dube in using the double machine learning method to estimate the elasticity of job posting durations with respect to the posting's advertised salary. The elasticities I find are small, around -0.07, but precisely estimated. Considering that the traditional model of the labor market would assume this elasticity to be large for a competitive market, this evidence indicates that some level of monopsony power exists in markets for remote labor.

Second, I examine the relationship between fully remote work and geography. It seems reasonable to assume that remote work is divorced from geography – that is, because the employee can work anywhere, the physical location of the employer would be irrelevant. If this were true, then the level of monopsony power in remote work across various commuting zones should not vary systematically from the national estimates provided previously. Empirically, this is not the case. Much like other papers have found when examining traditional labor markets (Azar et al. 2019), I find remote jobs based out of more populous commuting zones to have greater labor supply elasticity. At present, it is not clear why this is, but possible explanations could include differences in industry composition across the distribution of commuting zone population or better advertising by firms in large markets.

The results may seem discouraging to the prospect of remote work as a competitive, nationwide market with the potential to alleviate spatial wage inequality. However, this research at present does not estimate monopsony power for non-remote positions, which could be an important comparison group for drawing policy inference. As an example, the nationwide market for remote labor in a given occupation may still be more competitive than that occupation's local labor markets for in-person labor, especially if those areas are small. More work is needed to establish the role of remote work in shaping spatial wage inequality.

APPENDIX A. Additional Material for "Skills, Matching, and Skill Specificity Across Space"

Table A1. Five Most and Least Specific Skills According to Closeness Centrality

Most Specific	Least Specific
Production Process	Go-to-Market Strategy
Civil Engineering	Revenue Growth
Substance Abuse	Presentations
Ajax	Sales Strategy
Microbiology	Sales Forecasting

Table A2. Most Similar Skills to "Stata" for Euclidean Distance Versus Cosine Similarity

Euclidean Distance		Cosine Similarity	
Skill		Skill	
Scientific Studies	24.83	Literature Reviews	0.80
Analytic Applications	25.26	Survey Development	0.70
Grant Applications	27.45	Questionnaire Design	0.69
Syntax	28.25	Econometric Modeling	0.67
Information Synthesis	28.88	Empirical Research	0.66
Predictive Modeling	29.08	Applied Statistics	0.65
Statistical Software	29.46	Decision Tree Learning	0.63
Performance Analysis	29.50	Exploratory Data Analysis	0.63
Data Synthesis	29.61	Causal Inference	0.63
Quantitative Data Analysis	29.82	Multivariate Analysis	0.62

Table A3. Examples of Skills in Job Posting and Matched Resume

Skills in Job	Skills in Matched Resume
['Communications', 'Customer Service', 'Problem Solving', 'Computer Literacy', 'Scheduling', 'Mathematics', 'Prioritization', 'Warehousing', 'Quick Learning', 'Listening Skills', 'Setting Appointments']	['Customer Service', 'Detail Oriented', 'Lifting Ability', 'Merchandising', 'Time Management', 'Loading And Unloading', 'Retail Sales', 'Cash Handling', 'Inventory Control', 'Shipping And Receiving', 'Cashiering']
['Communications', 'Customer Service', 'Sales', 'Operations', 'Problem Solving', 'Lifting Ability', 'Mathematics', 'Warehousing', 'Good Driving Record', 'Forklift Truck', 'Safety Standards', 'Shipping And Receiving', 'CDL Class B License']	['Warehousing', 'Purchasing', 'Inventory Management', 'Supply Chain', 'Operations Management', 'Inventory Control', 'Pallet Jacks', 'Cycle Counting', 'Supply Chain Management', 'Forklift Operation', 'Order Picking']
['Communications', 'Customer Service', 'Sales', 'Leadership', 'Computer Literacy', 'Merchandising', 'Scheduling', 'Restaurant Operation', 'Cleanliness', 'Energetic', 'Quality Control', 'Food Services', 'Team Oriented', 'Operational Excellence', 'Leadership Development', 'ServSafe Certification']	['Customer Service', 'Sanitation', 'Inventory Management', 'Food Services', 'Food Safety And Sanitation', 'Greeting Customers', 'Cooking', 'Certified Nursing Assistant', 'Food Preparation', 'Restaurant Management', 'Cashiering', 'ServSafe Certification']
['Communications', 'Leadership', 'Detail Oriented', 'Problem Solving', 'Rehabilitation', 'Tactfulness', 'Collaboration', 'Treatment Planning', 'Geriatrics', 'Patient Assistance', 'Long-Term Care', 'Speech-Language Pathology', 'Orthopedics', 'Nursing Homes', 'Neurology']	['Nursing', 'Basic Life Support (BLS) Certification', 'Cardiopulmonary Resuscitation (CPR)', 'Medical Records', 'Advanced Cardiovascular Life Support (ACLS) Certification', 'Electronic Medical Record', 'Pediatric Advanced Life Support', 'Emergency Medicine', 'Critical Care Registered Nurse (CCRN)', 'Health Administration', 'Health Education']
['Communications', 'Customer Service', 'Operations', 'Professionalism', 'Billing', 'Workflow Management', 'Customer Relationship Management', 'Collections', 'Balancing (Ledger/Billing)', 'Medical Terminology', 'Internal Controls', 'Patient Assistance', 'Month-End Closing', 'Medical Billing And Coding']	['Auditing', 'Data Entry', 'Filing', 'Spreadsheets', 'Telephone Skills', 'Accounts Receivable', 'Accounts Payable', 'Deposit Accounts', 'Bookkeeping', 'Lawsuits', 'QuickBooks (Accounting Software)', 'Tax Returns']

Table A4. Comparison of Instrument Relevance

	First-stage F	Anderson-Rubin Wald F
1920 Population	1529.47	1512.55
1870 Population	1172.74	1165.5
Sedimentary Bedrock	30.81	31.61
Seismic Hazard	129.85	128.04
Landslide Hazard	93.68	92.13

Note: Measures come from regression of skill-level mismatch measure on instrumented log of population with no interaction effects included.

Table A5. Sensitivity of Results to Instrument Choice

	(1)	(2)	(3)	(4)	(5)	(6)
log(pop$_{c,2021}$)	-0.109***	-0.056***	0.069***	-0.068***	-0.107***	-0.054***
	(0.0008)	(0.0016)	(0.1256)	(0.0061)	(0.1210)	(0.0015)
1(high$_s$) × *log(pop$_{c,2021}$)*	-0.004***	-0.004***	-0.007***	-0.005***	-0.005***	-0.005***
	(0.0002)	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0002)
1(low$_s$) × *log(pop$_{c,2021}$)*	0.048***	0.038***	0.055***	0.049***	0.048***	0.046***
	(0.0002)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0002)
Constant	1.55***	0.916***	-0.692***	1.042***	1.535***	0.858***
	(0.0097)	(0.0206)	(0.1579)	(0.0775)	(0.1521)	(0.0188)
Instruments						
1920 Population	✓					✓
1870 Population		✓				✓
Sedimentary Bedrock			✓			✓
Seismic Hazard				✓		✓
Landslide Hazard					✓	✓
N	1,309,156	1,047,605	1,325,622	1,325,622	1,325,622	1,047,605

Note: Results are for the 3,000 most common skills. The outcome is the absolute value of the difference in the percentage of postings and resumes listing a skill.

APPENDIX B. Additional Material for "Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia"

Further discussion of non-winner decisions

Table B1 illustrates the fact that students who lose an enrollment lottery spend about half as many days enrolled in GA Pre-K compared to the average student who was never on any waitlist. Students who were never on a waitlist spend 85.9% of days between August 15th and May 31st enrolled in GA Pre-K, whereas students who do not win the lottery spend only 44.2% of the same time period enrolled in GA Pre-K. For non-winners, we fail to observe them in any GA Pre-K program for 55.8% of days in that span, indicating that they were either not attending any pre-K, or attending a pre-K program not administered by GA Pre-K. Children who do not win the lottery but attend GA Pre-K tend to spend the most time enrolled in non-SBPK sites, followed by SBPK sites other than their lottery school.

Table B1. Average Number of Days Spent in Each Type of Pre-K Site Between August 15th and May 31st

	Non-winners		Never on Waitlist	
	Mean	% of Days	Mean	% of Days
Days Not Enrolled in Pre-K	160.22	55.8	40.51	14.1
Days Enrolled in GA Pre-K	126.63	44.2	246.23	85.9
Days in Non-SBPK	73.96	25.6	140.85	48.7
Days in Any SBPK	52.67	18.2	105.37	36.5
Days in Preferred SBPK	32.52	11.3	-	-
Days in Other SBPK	20.15	7.0	-	-

Note: If a student loses a lottery for an SBPK, that SBPK is considered "preferred" by that student. The time period is 289 days between August 15th and May 31st.

APPENDIX C. Additional Material for "Monopsony in the Market for Remote Work"

Figure C1. Distribution of Calculated Job Posting Durations

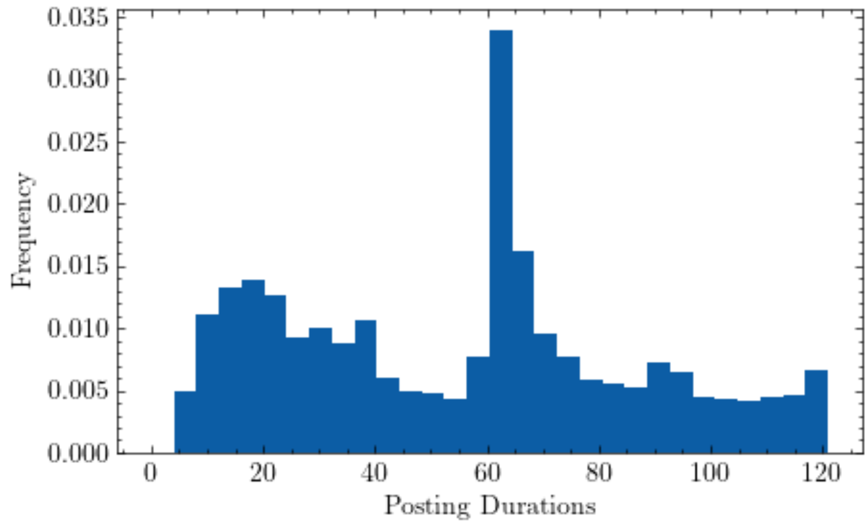


Figure C2. Distribution of Salary Ranges as Proportion of Salary Midpoint

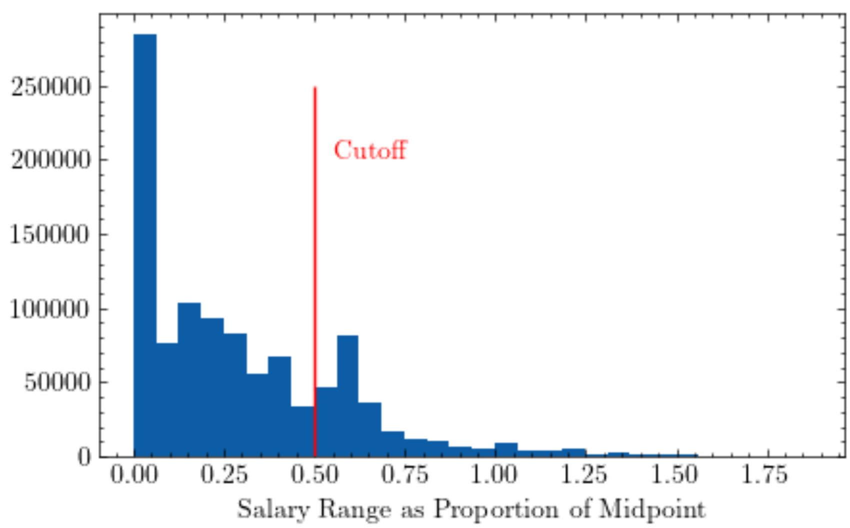
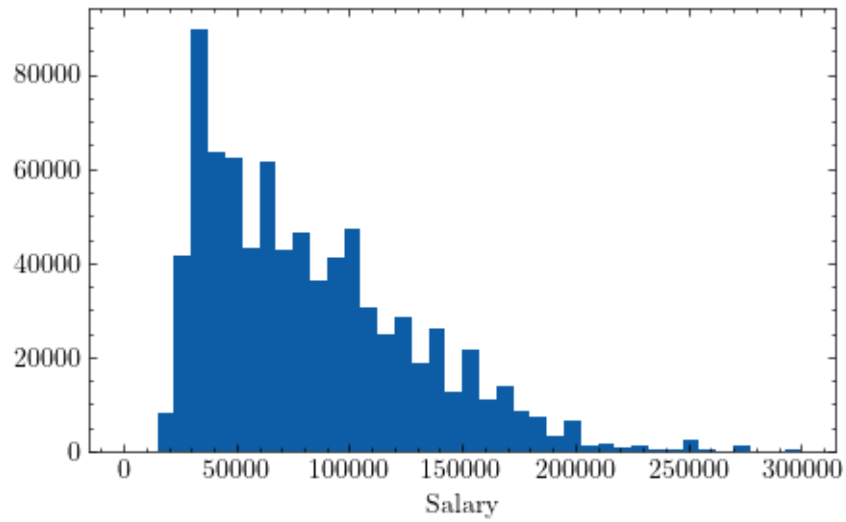


Figure C3. Distribution of Advertised Salaries in Remote Job Postings



REFERENCES

Abel, J., & Richard D. (2015). Agglomeration and job matching among college graduates. *Regional Science and Urban Economics* Vol. 51, No. 14.

Andersson, F., Burgess, S., & Lane, J. (2007). Cities, matching, and the productivity gains of agglomeration. *Journal of Urban Economics* Vol. 61, No. 112.

Atalay, E., Sotelo, S., & Tannenbaum, D. (2021). The Geography of Job Tasks. *Federal Reserve Bank of Philadelphia Working Paper Series WP21-27.*

Arnold Ventures (2017). Evidence Summary for the Abecedarian Project. Available at evidencebasedprograms.org/document/abecedarian-project-evidencesummary

Austin, B., Glaeser, E., & Summer, L. (2018). Saving the heartland: Place-based policies in 21st Century America. *Brookings Papers on Economic Activity Conference Drafts, March 8-9, 2018.*

Autor, D. (2019). Work of the past, work of the future. *Paper presented at the Richard T. Ely Lecture from American Economics Association Annual Meetings, Atlanta, GA*

Autor, D., Levy, F., and Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279-1333.

Azar, J., Marinescu, I., & Steinbaum M. (2019). Measuring Labor Market Power Two Ways. *AEA Papers and Proceedings,* 109, 317-321.

Azar, J., Marinescu, I., Steinbaum, M., & Taska, B. (2020). Concentration in US Labor Markets: Evidence from Online Vacancy Data. *Labour Economics*, *66*

Bacolod, M., Blum, B. S., & Strange, W. C. (2009a). Skills in the city. *Journal of Urban Economics*, 65(2), 136-153.

Bacolod, M., Blum, B. S., & Strange, W. C.(2009b). Urban interactions: Soft skills vs. specialization. *Journal of Economic Geography,* 9, 227-262.

Barnett, W. S. & Masse L. N. (2007). Comparative Benefit-Cost Analysis of the Abecedarian Program and its Policy Implications. *Economics of Education Review, 26*(1), 113-125

Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B. H., & Mihaylov E. (2022). The Shift to Remote Work Lessens Wage-Growth Pressures. NBER Working Paper 30197.

Bartik, T. J., Gormley, W., & Adelstein, S. (2012). Earnings benefits of Tulsa's pre-K program for different income groups. *Economics of Education Review*, *31,* pp. 1143-1161.

Baum-Snow, N. & Pavan, R., (2013). Inequality and City Size, *The Review of Economics and Statistics*, Vol. 95, No. 5

Baum-Snow, N., Freedman, M., & Pavan, R. (2018). Why Has Urban Inequality Increased? *American Economic Journal: Applied Economics*. Vol. 10.

Belfield, C. (2006). The Fiscal Impacts of Universal Pre-K: Case Study Analysis for Three States. *Invest in Kids Working Paper No. 6.*

Belfield, C., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup. *The Journal of Human Resources*, *41*(1), pp. 162-190.

Boal, W. M. & Ransom M. R. (1997). Monopsony in the Labor Market. *Journal of Economic Literature, 35*, pp. 86-112.

Card, D. & Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, *84*(4), pp. 772-793

Carneiro, P. M. & Heckman, J. (2003). Human Capital Policy. *IZA Discussion Paper* No. 821.

Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *American Economic Review* Vol. 86, No. 1.

Chen, Y., Cortes, P., Kosar, G., Pan, J., & Zafar, B. (2023). The Impact of COVID-19 on Workers' Expectations and Preferences for Remote Work. *AEA Papers and Proceedings*, *113*, pp. 556-561.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1)

Chor, E., Andresen, M. E., & Kalil, A. (2016). The impact of universal prekindergarten on family behavior and child outcomes. *Economics of Education Review*, *55*, pp. 168-181.

Combes, P.P., Duranton, G., Gobillon, L., & Roux, S. (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. *Agglomeration Economics*. Edited by Edward Glaeser.

Combes, P.P. & Gobillon, L. (2015). The Empirics of Agglomeration Economies. *Handbook of Regional and Urban Economics*. Edited by Gilles Duranton, J. Vernon Henderson, & William C. Strange. Vol. 5.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31-47.

Cunha, F., & Heckman, J. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, 7(2-3), 320-364.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, *1*, 697-812.

Currie, J. (2001). Early Childhood Education Programs. *Journal of Economic Perspectives*. *15*(2), pp. 213-238.

Currie, J & Thomas, D. (1995). Does Head Start Make a Difference? *The American Economic Review*, *85*(3), pp. 341-364.

Dauth, W. Findeisen, S. Moretti, E. & Suedekum, J. (2016). Spatial wage disparities - Workers, firms, and assortative matching. *Working Paper.*

Dauth, W., Findeisen, S., Moretti, E., & Suedekum, J. (2022). Matching in Cities. *Journal of the European Economic Association Vol 20, No. 4*.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), pp. 111-134.

Di Addario, S. (2011). Job Search in Thick Markets. *Journal of Urban Markets* Vol. 69, No. 3.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3), 479-524.

Dube, A., Jacobs, J., Naidu, S., & Suri, S. (2020). Monopsony in Online Labor Markets. *AER: Insights*, *2*(1), pp. 33-46.

Duranton, G. & Puga, D. (2003). Micro-foundations of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics*

Durkin, K., Lipsey, M. W., Farran, D. C., & Wiesen, S. E. (2022). Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade. *Developmental Psychology*, *58*(3), 470.

Florida, R., Mellander, C., Stolarick, K., & Ross, A. (2012). Cities, skills, and wages. *Journal of Economic Geography,* 12(2), 355-377

Fredriksson, P., Hensvik, L., & Skans, O. N. (2018). Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility. *American Economic Review* Vol. 108, No. 11

Friedman-Krauss, A. H., Barnett, W. S., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., Gardiner, B. A., & Jost, T. M. (2022). The State of Preschool 2021: State Preschool Yearbook. *National Institute for Early Education Research*.

García, J. L., Heckman, J. J., Leaf, D. E., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, *128*(7), 2502-2541.

Giannone, E. (2018). Skill-biased technical change and regional convergence. Working Paper.

Goldhaber, D., & Özek, U. (2019). How much should we rely on student test achievement as a measure of success? *Educational Researcher*, *48*(7), 479-483.

Goldring, T. (2020). Children in Georgia's Pre-K Program: Characteristics and Trends, 2011-12 to 2018-19. Georgia Policy Labs.

Guvenen, F., Kuruscu, B. Tanaka, S., & Wiczer, D. (2020). Multidimensional Skill Mismatch. *American Economic Journal: Macroeconomics*. Vol. 12, No. 1.

Harmon, N. (2013). Are Workers Better Matched in Large Labor Markets? *Job Market Paper*.

He, H., Neumark, D., & Weng, Q. (2021). Do Workers Value Flexible Jobs? A Field Experiment *Journal of Labor Economics*, *39*(3)

Heckman, J. (2000). Policies to foster human capital. *Research in Economics*, *54*(1), 3-56.

Heckman, J. (2008). Schools, skills, and synapses. *Economic Inquiry*, *46*(3), pp. 289-324.

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010a). The rate of return to the HighScope Perry Preschool Program. *Journal of public Economics*, *94*(1-2), 114-128.

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010b). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative economics*, *1*(1), 1-46.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052-2086.

Heckman, J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411-482.

van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, *66*, pp. 206-222.

Ingram, B. F. & Neumann, G.R. (2006). The returns to skill. *Labour Economics*, 13(1), 35-59.

Justice, L. M., Bowles, R. P., Pence Turnbull, K. L., & Skibbe, L. E. (2009). School readiness among children with varying histories of language difficulties. *Developmental Psychology*, *66*(2), 460.

Karoly, L. A. (2016). The Economic Returns to Early Childhood Education. *The Future of Children. 26*(2), pp. 37-55.

Kingsley, S. C., Gray, M. L., & Suri, S. (2015). Accounting for Market Frictions and Power Asymmetries in Online Labor Markets. *Policy & Internet, 7*(4)

Lee, V. E., Brooks-Gunn, J., Schnur, E., & Liaw, F. R. (1990). Are Head Start Effects Sustained? A Longitudinal Follow-Up Comparison of Disadvantaged Children Attending Head Start, No Preschool, and Other Preschool Programs. *Child Development*, *61*(2), pp 495 – 507

Lipsey, M. W., Farran D. C., & Durkin. K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly, 45*, 155-176

Luccioletti, C. (2023). Labor Market Power Across Cities. Working Paper.

Ludwig, J. & Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *The Quarterly Journal of Economics*.

Manning, A. (2003) *Monopsony in Motion: Imperfect Competition in the Labor Markets.* Princeton University Press.

Naidu, S., Posner, E. A., & Weyl, G. (2018). Antitrust Remedies for Labor Market Power. *Harvard Law Review*, *132*, pp. 536-601.

Peisner-Feinberg, E., LaForett, D., Schaaf, J., & Hildebrandt, L. (2013). Local Variations in Enrollment Processes in Georgia's Pre-K Program: Findings from the 2012-2013 Evaluation Study. Georgia's Pre-K Program Evaluation Project.

Petrongolo, B. & Pissarides C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature* Vol. 39 No. 2. pp. 390-431.

Poletaev, M. & Robinson, C. (2008). Human capital specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984 – 2000. *Journal of Labor Economics*, 26(3), 387-420.

Robinson, J. (1969). The Economics of Imperfect Competition. *Palgrave Macmillan London.*

Schmutz, B., Sidibé, M., & Vidal-Naquet, É. (2021). Why Are Low-Skilled Workers Less Mobile? The Role of Mobility Costs and Spatial Frictions. *Annals of Economics and Statistics*. No. 142, pp. 283-304.

Stipek, D. (2006). No child left behind comes to preschool. *The Elementary School Journal*, *106*(5), 455-466.

Tsur, S., Brin, S., Motwani, R., & Ullman, J. D. (1997). Dynamic Itemset Counting and Implication Rules for Market Basket Data. *ACM SIGMOD Record*. Vol. 26, No. 2. pp 255–264.

Watts, T. W. (2020). Academic achievement and economic attainment: Reexamining associations between test scores and long-run earnings. *Aera Open*, *6*(2), 2332858420928985.

Weinstein, A. & Patrick, C. (2020). Recession-proof skills, Cities, and Resilience in Economic Downturns. *Journal of Regional Science* 60(2): 348-373.

Wheeler, C. (2008). Local market scale and the pattern of job changes among young men. *Regional Science and Urban Economics* Vol. 38, No. 2.

Wolff, E.N. (2003). Skills and changing comparative advantage. *Review of Economics and Statistics*. 85(1), 77-93.

Wooldridge, J. M. (2001). Econometric Analysis of Cross Section and Panel Data, *MIT Press.*

Zlatko, N. & Ajwad, M.I. (2014). Cognitive and non-cognitive skills affect employment outcomes: evidence from Central Asia. *The World Bank*, Washington, USA.

VITA

Henry Woodyard is an applied microeconomist with an interest in topics at the intersection of labor and urban economics. After receiving his bachelor's degree in economics from the University of Southern Mississippi in 2018, he began work that year toward his doctorate in economics at Georgia State University. While at Georgia State, Henry also worked with the Georgia Policy Labs (GPL) in the Metropolitan Atlanta Policy Lab for Education (MAPLE). There, he worked on various topics including, but not limited to, virtual instruction, learning loss during the pandemic, early-childhood education, student behavior, and academic mobility. Henry has also received several awards during his time at Georgia State. These include the *University Fellowship*, *Outstanding Graduate Research Assistant Award,* and *The AYSPS Excellence in Teaching Economics Award.*