DISSERTATION SERIES 2026:2

Networks, Norms, and Non-market Subsidy

Mohammad Zariab Hossain

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2 (Hall 2), Kyrkogårdsgatan 10, Uppsala, Wednesday, 10 June 2026 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Research Director Hanna Pesola (VATT Institute for Economic Research).

ISSN 1651-4149

Economic Studies 236

Mohammad Zariab Hossain

Networks, Norms, and Non-market Subsidy

Department of Economics, Uppsala University

Visiting address: Kyrkogårdsgatan 10, Uppsala, Sweden
Postal address: Box 513, SE-751 20 Uppsala, Sweden
Telephone: +46 18 471 00 00
Telefax: +46 18 471 14 78
Internet: http://www.nek.uu.se/

ECONOMICS AT UPPSALA UNIVERSITY

The Department of Economics at Uppsala University has a long history. The first chair in Economics in the Nordic countries was instituted at Uppsala University in 1741.

The main focus of research at the department has varied over the years but has typically been oriented towards policy-relevant applied economics, including both theoretical and empirical studies. The currently most active areas of research can be grouped into six categories:

* Labour economics
* Public economics
* Macroeconomics
* Microeconometrics
* Environmental economics
* Housing and urban economics

Mohammad Zariab Hossain

Networks, Norms, and Non-market Subsidy

Three essays on Public Policy



UPPSALA
UNIVERSITET

Uppsala 2026

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2 (Hall 2), Kyrkogårdsgatan 10, Uppsala, Wednesday, 10 June 2026 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Research Director Hanna Pesola (VATT Institute for Economic Research).

Abstract
Hossain, M. Z. 2026. Networks, Norms, and Non-market Subsidy. Three essays on Public Policy. *Economic studies* 236. 140 pp. Uppsala: Department of Economics, Uppsala University. ISBN 978-91-506-3177-7.

Essay I: This paper provides the first causal evidence of peer-driven behavioral spillovers in public employment services (PES). Using rich Swedish administrative data from 2004–2018 and a difference-in-differences design exploiting caseworker moves across local PES offices, I show that caseworker's broader vacancy referral behavior (i.e., vacancy referrals outside a job seeker's stated preferences) responds strongly to the local practice environment, particularly peer norms. Moving to an office with a 1 pp higher average broader-referral rate increases a caseworker's own broader-referral propensity by 0.35 pp. I then examine how these practice-environment shocks of the movers (caseworkers) affect their job seekers; and I find that job seekers' re-employment probability (within 180 days) reduces by 1.5 pp when their caseworkers are exposed to 10 pp higher broader referral environment. These findings are consistent with evidence from mandated broader-search policies (Van der Klaauw and Vethaak, 2022) and stand in contrast to settings where broader referrals are given as informational nudge (Belot et al., 2019, 2022). Taken together, the evidence shows that local peer norms are important and have consequences for job seekers' reemployment prospects.
Essay II: This paper studies whether public employment services can mitigate occupational gender segregation by shaping job seekers' exposure to gender-atypical employment opportunities. Using rich administrative data from Sweden and quasi-random assignment of job seekers to caseworkers based on date-of-birth rules, we causally identify the effects of caseworkers' gender non-stereotypicalness (with regard to vacancy referral) on job seekers' labor market outcomes. We show that exposure to more gender non-stereotypical caseworkers significantly increases reemployment probabilities and leads to persistent earnings gains lasting up to five years. Our findings highlight the role of labor market intermediaries in shaping occupational sorting and demonstrate that norm-challenging job-search assistance may improve the efficiency in the labor market.
Essay III: This paper provides the causal evaluation of a large-scale education subsidy program in Bangladesh that finances teacher salaries in non-government secondary schools. Using nationally representative data from 2012-2023 and a regression discontinuity design, it examines whether government-financed salary support improved student achievement and school participation. The results reveal no significant effects on passing rates, share of distinctions in public examinations, or teacher composition, suggesting that fiscal relief alone did not translate into improved learning outcomes. While the estimates indicate a modest increase in girls' access to secondary schooling, the evidence remains too limited to support a strong conclusion. In addition, school outcomes appear broadly similar across regions with differing levels of poverty and educational corruption. The reform's broader impact may have been constrained by the absence of changes in teaching personnel and, consequently, class sizes, as well as by weak accountability, limited performance incentives, and documented dissatisfaction among MPO teachers.

Keywords: job seekers, labor market policy, caseworkers, broader vacancy referrals, peer effects, occupational segregation, gender-norms, non-stereotypical vacancy referrals, secondary education, payroll subsidy

Mohammad Zariab Hossain, Department of Economics, Box 513, Uppsala University, SE-75120 Uppsala, Sweden.

© Mohammad Zariab Hossain 2026

This publication is distributed under the terms of the Creative Commons Attribution Non-Commercial No-Derivatives (CC BY NC ND) license.

ISSN 0283-7668
ISBN 978-91-506-3177-7
URN urn:nbn:se:uu:diva-584671 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-584671)

To my family

Acknowledgements

First of all, I can't thank the Almighty Creator enough for this entire journey! It was a hefty challenge for a man in his late thirties to dive into the tumultuous PhD life along with his wife and a one-and-a-half-year-old toddler, just as the whole world was being thrown into COVID-19 chaos (which also took away some beautiful souls that I cared for!). I still wonder how I got the courage to travel across a continent to finally land in Uppsala, which I will continue to cherish for the rest of my life!

Writing these acknowledgements feels both joyful and overwhelming—as if gratitude itself has too many references to organize. So many people made this journey possible, and I owe each of them more than I can properly express. First of all, I would like to sincerely thank Mikael Bask and the PhD Admission Committee at Uppsala for taking a chance on me and welcoming me into this prestigious department, where I had the privilege of learning from scholars at the forefront of economic research. I also owe Bask a special apology for withdrawing my 2020 admission at the very last moment because of COVID-related difficulties—I am quite sure that did not make me the easiest applicant to deal with. Yet even then, I had already begun to feel that I belonged at the Uppsala Economics Department. So being admitted in 2021 meant a great deal to me; it felt like Uppsala kindly decided not to give up on me just yet.

Johan! I came to this department planning to specialize in development economics, but the impression you left on me while teaching econometrics was something special; I felt an immediate connection. If I had to choose my supervisor again, I would still choose you again and again. In the highly convincing disguise of an academic supervisor, you have been a true friend: always asking about my family, believing in me even when things were difficult, and supporting me in every possible way. Without you, this thesis would simply not exist—and that is hardly an overstatement. Writing this paragraph already stirs a mix of discomfort and emotion, knowing that our supervisor–supervisee journey is drawing to a close. But I will always miss those candid moments—your reverberating laughter in the dining room over lunches and fika, always full of positive energy.

Peter! When I applied to this PhD program in Uppsala, I had already hoped that you might become my supervisor—I even made sure to mention it in my motivation letter, just in case destiny was reading carefully. To me, you are a bit of an academic celebrity. Before coming to Uppsala, I had never met or spoken to an economist of your stature, so you can imagine how thrilled I was when you became my co-supervisor. I am truly grateful for your thoughtful feedback on my thesis chapters.

Jonas! You are a true mentor, and I cannot thank Johan enough for pairing me with you. You taught me everything from Swedish registry data to coding in Stata, for discussions related to econometrics and what not, and somehow managed to survive my endless stream of emails and questions without ever filing a formal complaint. I am still convinced that your email response time defies basic human limitations. I will also always remember our lively discussions in front of the whiteboards, which helped me enormously.

When I started my PhD, I thought it would be a solitary journey. Clearly, I had misunderstood the syllabus. There is an old adage that says, "We were never meant to carry our burdens alone," and I truly feel blessed to have had the nicest cohort in the world. I learned so much from you all, and you will always remain in my thoughts. Adam, Chris, Henning, Jakub, and Tram (listed in alphabetical order, so no one can start a side argument): words can hardly do justice to how much you mean to me. Our lunches, weekend get-togethers, Sirius games, and afternoon walks through the cemetery—have given me memories that will stay with me long after this chapter ends. You are more than friends to me; you are truly a family within a family.

No acknowledgment would be complete without thanking my Uppsala gurus, from whom I have been deeply enlightened. Once a pundit said, "A teacher affects eternity, but can never tell where the influence stops." I feel immensely grateful to have had Mikael Bask, Chuan-Zhong Li, Torben Mideksa, Johan Vikström, Peter Fredriksson, Erik Öberg, Oskar Nordström Skans, Ola Andersson, Georg Graetz, and Luca Repetto as my teachers on this journey. I believe your teachings will continue to influence me for many years to come. I also owe a great deal to the Uppsala Labor Group (aka ULG) for their invaluable feedback on my thesis, especially Oskar, Georg, Lena, Yoko, Helena, Olof, and Eva. I would also like to thank Hiep, Erik, Majken, Sebastian, Elisavet, Xuemeng, Suna, Gilson, Christian, Ricard, Svante, Olle, Dogan, Simon, Erika, Rinni, Jakob, Hanfeng, Tianze, Amin, Maria Elena, Caio, and Kristina for the many wonderful discussions we shared on different occasions. My heartfelt thanks also go to the administrative staff at NEK, especially Ulrika, for always being there whenever I needed help and for saving me repeatedly from bureaucratic adventures.

During my second year, Johan arranged for me to have a second office at IFAU, which very quickly became my second home in Uppsala. Over the course of four years, I met many wonderful and kind-hearted people there: Åsa, Maria, Melica, Martin S, Ali, Malin L, Malin TA, Björn, Anna, Anahid, Sara M, Sara R, Caroline, Anders, Adrian, Olof, Spencer, Klara, Andrei, Lisa, Erica, Martin L, Martin N, Ulrika, Mattias, Josefin, Ingeborg, Kristina, Håkan, Jan, Linus, Patrik, Jörgen, and Helge. I greatly appreciate Patrik for your kind words and care, Jörgen for your patience with my endless IT-related questions; and Olof for your valuable feedback during my final seminar. My heart will always hold on to the memories made in the IFAU PhD room. Henning, I truly loved your company at IFAU, and I must admit that it would have been a

completely different world without you there. So many of my IFAU memories revolve around our mini-break discussions, lunch walks to ICA, and many other small moments that made the place special. A big thank-you to Olle, Oona, and Gabriel for your wonderful presence during my time at IFAU and for the backyard floorball.

Before the doctoral study, I had the good fortune to meet several inspiring people who nurtured my thirst for economics at different stages of my journey. I am deeply indebted to Gour Gobinda Goswami, Professor of Economics at North South University, Bangladesh, whose words of encouragement and motivation fanned into flame my hunger for knowing more. I also gratefully acknowledge the contribution and support of Late Habibur Rahman (UIU), Md. Abdul Matin (DU), Md. Ashraf Uddin (SUST), Ada Ferrer-i-Carbonell (UAB), Enriqueta Aragonès (UAB), Pietro Dindo (UVenice), and Giacomo De Giorgi (UGeneva) for their teachings at different points in time.

To speak honestly of this journey, I must begin with those who made me who I am today. My father was always deeply hopeful about my future, but sadly, he left this world when I was only twelve. Abba, I know you would have been so happy to see this feat, and I like to believe you are smiling from the other side—probably with the calm satisfaction of someone who always knew I would make it. Amma, you are the purest example of courage I have ever known. Widowed in your early thirties, without work and with every reason to collapse under the weight of life, you instead carried us with unimaginable strength. You never once put yourself first, and your selfless devotion and countless sacrifices for us will remain with me until my last breath. I do not have words big enough for what you mean to me, and you are the reason for who I am today. Urmi, my only sister, thank you for the beautiful childhood we shared together. Even though we spent parts of it behaving less like siblings and more like two rivals fighting over absolutely nothing, you have always been one of the reasons behind my ambitions. Your faith in me—that I would somehow conquer the world—has stayed with me more than you know. And Aiza, my sweet niece, you are the cutest little person who brought so much laughter into my life, especially during time of utter distress.

I also remain forever indebted to my maternal uncles—Anjum Mama, Asif Mama, and Ahsan Mama—and my aunts—Boro Aunty, Mejho Aunty, Acchu Aunty, Hasnat Aunty, and Ishrat Aunty—for their unwavering support in the absence of my father. Some debts are simply too large to be repaid in this lifetime. I also gratefully acknowledge the limitless trust of my cousins, especially Russell bhaiya, for helping me when I truly needed it most. My deepest regards also go to my mother-in-law and father-in-law for their trust, care, and kindness, which have meant a great deal to me.

One of life's greatest blessings is to have friends whose wisdom and companionship shape who we become. I feel truly blessed to have had such friends in my life. My childhood friends—Shanto, Shuvro, Manik, Himu, Mona, Sukto, and Tomal—have contributed in no small measure to the person I have become

today. I am deeply grateful for your friendship, and I know that my accomplishments today are shaped in many ways by the thoughts, dreams, and moments we shared across different stages of life. I am also fortunate to have friends like Hasan and Mamun bhai, whose warmth and positive energy have added so much to my journey. I am equally grateful to my colleagues—Md. Mahbub Zaman Khan, Shimul Sen, Md. Mahbubul Alam Siddiqui, Dr. Shamsul Alam, Md. Zakir Hossen, Nawsher Ahmed Shikdar, MM Ashraf Reza Faridy, Shereen Akhter, Md. Mojibur Rahman and Md. Rezwanul Hoque—for their generous support during difficult times. Your kindness and readiness to help will remain with me long after this chapter closes.

Bithi! If this journey rests on any foundation at all, it is you. You have been the quiet force beneath everything, the unseen strength. In every season of this PhD, you were my shelter in chaos, my peace in exhaustion, and my courage when my own began to fail me. You gave me the freedom to pursue this dream by bearing more than anyone should ever have to bear alone. Whatever I have achieved, whatever dignity this thesis may hold, it carries your sacrifice in every line. No language I know is large enough to repay what I owe you. And finally Zavian, my precious boy! You are the purest light I have ever known. In the darkest and most demanding days of this journey, you arrived like joy itself, turning exhaustion into laughter and heaviness into wonder. After the longest days, your hugs and kisses could restore parts of me that I did not even know were broken. Your presence filled this life with warmth, meaning, and a kind of happiness that no achievement could ever rival. Even your endless questions—so sincere, so relentless, so beautifully impossible to escape—became reminders of how miraculous love can be in its smallest, most ordinary forms. If this thesis marks the end of a long chapter of my life, it is you two who gave that chapter its soul. Whatever these years took from me, you gave back more than I ever deserved: a home, a heart, and a reason to keep going.

Uppsala, April 2026
Mohammad Zariab Hossain

Contents

Introduction 13
References 22

Essay I: Vacancy Referrals and Broader Job Search: The Role of Caseworkers and Peer Spillovers 25
Introduction 26
Swedish Public Employment Service offices and caseworkers 30
Caseworker movement within Swedish PES 30
Broader referral behavior 31
Data and sample 32
Practice-environment based on broader referral 34
Practice-environment 34
Econometric framework and identification 36
Spillover effects among migrant caseworkers 36
Robustness analysis 38
Decomposition of practice-environment effect 39
Estimating job seeker's outcomes 39
Do caseworkers change behavior upon changing offices 40
Heterogeneity in spillovers 44
Decomposing the spillover effects 45
Effects on Job Seeker Outcomes 46
Heterogeneity in job seeker's outcomes 49
Conclusion 51
References 53
Appendix 55

Essay 2: Returns to Breaking the Norm: Evidence from Job Referrals 59
Introduction 60
Background: Gender-segregated labor-markets 63
Occupational segregation in the US and Sweden 63
Segregated job-search 64
Institutional setting 66
Public employment service and caseworkers 66
Stereotypical and non-stereotypical caseworkers 66
Date-of-birth assignment 67
Data & Descriptives 68
Data 68

Descriptive statistics ... 69
Defining gender stereotypical referral behavior ... 69
Identification strategy ... 71
Empirical strategy ... 71
Assessing the instrument ... 72
Effects of a gender non-stereotypical caseworkers ... 75
Average effects ... 75
Differences between male and female job seekers ... 78
Non-stereotypicalness or other caseworker traits? ... 79
Non-traditional jobs, broader search and/or labor marker tightness? ... 79
Effects on wages and occupations ... 81
Female job seekers ... 81
Male job seekers ... 82
Conclusions ... 83
References ... 85
Appendix ... 87

Essay 3: Financing Secondary Education: Closing Gaps or Boosting Grades? ... 95
Introduction ... 96
Institutional background ... 99
Some basic facts about secondary education in Bangladesh ... 99
The Monthly Pay Order (MPO) intervention ... 101
The 2018 MPO reform and selection criteria ... 102
Is MPO important for schools? ... 103
Challenges of MPO intervention ... 105
Data and sample ... 105
Econometric framework and identification ... 107
Identification ... 107
Empirical framework ... 108
Treatment assignment, covariate balance and bunching ... 110
Treatment assignment based on running variables ... 111
Covariate balance ... 111
Bunching ... 113
Does MPO affect school outcomes? ... 115
Do regional socioeconomic conditions and corruption matter for school performance? ... 119
Mechanism ... 122
Conclusion ... 126
References ... 128
Appendix ... 131

Introduction

While public policy is often evaluated through the lenses of resources, incentives, and formal program design, it is important to keep in mind that policies do not operate mechanically. Their effectiveness depends on how they are implemented, how they are mediated through organizations and frontline actors, and how individuals respond within social and institutional environments, where decisions are shaped not only by incentives, but also by information, norms, and institutional constraints. This dissertation is motivated by that broader observation. It studies how economic outcomes are shaped by the interaction between policy design, policy delivery, and the social environments in which policies operate. More specifically, it examines three closely related themes. First, it studies how public employment services affect labor market outcomes through the behavior of caseworkers, rather than merely through the formal existence of job-search assistance. Second, it considers how gender norms embedded in institutional practice may influence occupational mobility and inequality. Third, it asks why substantial public investment in education may fail to improve core performance outcomes when implementation environments are weak. Across these themes, the dissertation advances a common argument: policy effects depend not only on what governments do on paper, but also on who implements policy, under what norms, and within which institutional constraints.

The dissertation is situated at the intersection of labor economics, public economics, and the economics of education. Its chapters cover both developed and developing country settings and use quasi-experimental methods to identify causal effects. Despite the different domains, the essays are linked by a shared concern with how institutions mediate behavior. In the labor market essays, the focus is on caseworkers in the Public Employment Service (PES), who guide unemployed job seekers and thereby shape their opportunity sets. In the education essay, the focus is on whether a major public subsidy program in Bangladesh translated additional resources into improved school outcomes. Together, the essays speak to a set of fundamental public policy questions: *why do some interventions succeed while others disappoint? How important are organizational behavior and social norms for policy delivery? And to what extent can public institutions alter the choices, trajectories, and welfare of the individuals they serve?*

Unemployment, occupational segregation, and the role of public employment services

Unemployment remains one of the most persistent and consequential challenges facing modern economies. For individuals, it is associated with lost earnings, skill depreciation, lower future employability, and psychological strain. For society, it implies forgone output, higher fiscal costs, and weaker aggregate demand. Job-search assistance (JSA)[1] is one of the most common labor market policy tools used to address this problem. Typically delivered through caseworkers in Public Employment Service (PES) offices, JSA aims to improve job finding and labor market matching by helping unemployed workers overcome information frictions and identify suitable vacancies. A substantial empirical literature shows that JSA can raise re-employment probabilities and shorten unemployment spells (Card et al., 2010; Weber and Hofer, 2003; Crépon et al., 2005; Van den Berg et al., 2014; Cheung et al., 2024).

Building on this evidence, policymakers have increasingly emphasized broader search strategies that encourage job seekers to look beyond their previous occupations or stated preferences. The idea is that unemployed workers may underestimate the feasibility or desirability of jobs outside their prior experience, and that a wider search horizon may improve job finding or match quality. Some empirical work supports this view, especially for disadvantaged groups (Belot et al., 2019, 2025), while other studies find weak or even negative effects, suggesting that broader referrals may dilute search effort or direct workers toward less suitable vacancies (Dhia et al., 2022; Van der Klaauw and Vethaak, 2022). This mixed evidence points to the importance of how such strategies are implemented in practice.

Alongside unemployment, labor market inequality is another fundamental concern, closely linked to occupational segregation. The uneven distribution of men and women across occupations remains a central feature of modern labor markets and a major source of gender inequality. Although female labor force participation rose substantially and some forms of segregation declined over the second half of the twentieth century, progress toward gender-balanced occupations has slowed markedly since the early 2000s (Blau and Kahn, 2017; Blau et al., 2013). While a conventional view explains occupational sorting through comparative advantage or heterogeneous preferences, an increasingly influential alternative emphasizes gender norms, stereotypes, and institutional frictions that constrain opportunity sets and distort labor market allocation. Consistent with this view, a growing literature shows that occupational aspirations, applications, and career choices are shaped by beliefs about gender-appropriate work, social expectations, and employer-side frictions (Barbulescu and Bidwell, 2013; Card et al., 2025; Cejka and Eagly, 1999; Del Carpio and Guadalupe, 2022; Sorkin, 2017; Vattuone, 2024). Experimental and quasi-experimental

[1] It includes vacancy referrals, counseling, training guidance, monitoring, job suggestions, and career planning.

studies further show that employers treat men and women differently even when productivity differences are negligible (Azmat and Petrongolo, 2014; Baert, 2018; Card et al., 2025; Kuhn and Shen, 2023).

Public agencies therefore have a direct stake in both unemployment and occupational segregation. PES offices are especially relevant because they intervene directly during job search. Caseworkers do more than monitor effort: they identify vacancies, suggest occupations, and shape which opportunities job seekers perceive as realistic. Referral practices are thus not neutral. They may encode assumptions about gender and work, with consequences for occupational mobility, earnings, and the reproduction of inequality. Evidence from related contexts suggests that referral and recommendation behavior often exhibits gender homophily, especially in strongly stereotyped settings (Beugnot and Peterlé, 2020; Hederos et al., 2025).

These issues lie at the core of this dissertation. Although there is extensive research on the effectiveness of job-search programs, much less is known about the people who deliver them. Standard job search models typically assume that workers choose from a well-defined set of opportunities (Mortensen and Pissarides, 1994; Pissarides, 2000), but in practice those opportunity sets are often structured by PES caseworkers. Caseworkers decide which vacancies to refer, how strongly to encourage broader search, and how to balance organizational objectives against individual needs. Their behavior is therefore likely to be an important determinant of policy effectiveness.

Two features of caseworker behavior are central to this dissertation. First, caseworkers operate within organizational environments and may adapt to the norms and practices of their peers. A large literature shows that behavior is shaped by peers and local social environments in workplaces and other institutions (Manski, 2000; Sacerdote, 2001; Bandiera et al., 2009; Mas and Moretti, 2009), yet this possibility has received little attention in the context of public employment services. If caseworkers adapt to office-specific norms, then formally identical JSA programs may be delivered differently across local settings, with consequences for job seekers' outcomes.

Second, caseworkers may differ in the gender-related norms they bring into their work. If they guide job seekers toward occupations in ways that reflect stereotypical perceptions, PES offices may reinforce existing inequalities rather than simply improve job finding. Conversely, if they expand job seekers' opportunity sets in non-stereotypical ways, PES may promote not only labor market matching but also social mobility and reduced occupational segregation. In this sense, caseworker heterogeneity is not merely administrative; it may shape the substance of policy itself.

These concerns motivate the first two essays of the dissertation. Both take seriously the idea that public policy operates through frontline actors whose own behavior is shaped by social and organizational environments. The contribution is therefore both substantive and conceptual: substantively, the essays identify the effects of peer environments and gender-stereotypical referral practices;

conceptually, they argue that labor market policy should be studied not only as a set of rules and resources, but also as a socially mediated process of implementation.

Public spending, institutional quality, and policy effectiveness in education

The third essay moves from labor market policy to education, but it is driven by a closely related concern. Governments around the world invest heavily in education because schooling is central to human capital accumulation, social mobility, and long-term economic development (Psacharopoulos, 1994; Hanushek, 2006). Yet there is long-standing disagreement in the literature about whether greater public spending in education necessarily improves outcomes. A broad lesson from this literature is that the effects of educational investment depend strongly on context. Identical or similar interventions can generate different results depending on local governance, bureaucratic capacity, accountability systems, and the distinction between expenditure that expands resources and expenditure that changes instructional quality (Hoxby, 2000; Dustmann et al., 2003; Marshall, 2009).

This issue is especially salient in settings where institutions are weak and monitoring is limited. In such environments, additional funding may improve access or financial sustainability without necessarily improving learning outcomes. Resources may fail to translate into changes in classroom practice, teacher effort, or pedagogical quality. These concerns motivate the third essay, which studies a major public intervention in Bangladesh's secondary education sector: the Monthly Pay Order (MPO) program. Under this system, the government finances a substantial portion of teacher payroll costs in non-government secondary and higher secondary institutions, particularly in rural areas. The program is both longstanding and quantitatively important, making it an ideal setting in which to study whether a large public subsidy improves educational outcomes.

The Bangladesh setting broadens the dissertation's scope in two ways. First, it extends the analysis from labor market policy in a high-income administrative context to education policy in a lower-income setting with different institutional constraints. Second, it sharpens the dissertation's general theme by showing that the problem of implementation is not limited to frontline discretion in public employment services. Even where policies are less individualized and more financial in nature, outcomes still depend on the institutional environment through which resources are translated into action.

In that sense, the education essay complements the labor market essays. The first two essays show that caseworker behavior, peer effects, and gender norms can shape the effectiveness of a policy tool that is widely considered beneficial. The third essay shows that even substantial and sustained public spending may

fail to improve intended outcomes if the implementation environment is weak. Together, the essays therefore challenge simple input-output views of policy and point instead to the centrality of institutions, norms, and organizational processes.

Essay 1: Vacancy referrals, broader job search, and peer spillovers

While vacancy referrals are a core instrument of job-search assistance, relatively little is known about the factors that shape caseworkers' referral behavior or whether that behavior has meaningful consequences for job seekers. Because referrals directly shape the opportunity sets individuals face, they are likely to affect both the speed and the quality of re-employment. The first essay, titled ***"Vacancy Referrals and Broader Job Search: The Role of Caseworkers and Peer Spillovers"***, studies how office-level norms, shaped by peers within the workplace, influence caseworkers' delivery of services to unemployed job seekers.

Using Swedish administrative data for 2004–2018, the essay follows caseworkers who move across local Public Employment Service offices and examines how their referral behavior changes after relocation. Exploiting plausibly exogenous variation generated by caseworker mobility, and using difference-in-differences and event-study designs, the analysis finds strong evidence that caseworkers adapt to the norms of their new office. A one percentage point increase in the destination office's broader referral rate increases the migrant caseworker's own broader referral behavior by 0.346 percentage points. This pattern suggests rapid alignment with local workplace norms, consistent with broader evidence on peer effects in organizational settings (Sacerdote, 2001; Bandiera et al., 2009). A decomposition further shows that peer influence accounts for most of the overall spillover effect: 0.284, compared with 0.062 for office-specific factors.

The essay then examines whether exposure to different practice environments affects the labor market outcomes of the job seekers served by these caseworkers. This question is especially relevant because the evidence on broader job search is mixed: some studies find positive effects (Belot et al., 2019), while others show that broader search requirements may be counterproductive (Van der Klaauw and Vethaak, 2022). Rather than studying direct policy mandates or informational nudges, the essay asks how exposure to a more intensive broader-referral environment, through changes in caseworkers' workplace settings, affects client outcomes.

The results show that when a migrant caseworker moves to an office with a 10 percentage point higher broader referral intensity, the probability that their clients exit unemployment within 180 days falls by 1.5 percentage points. One

possible mechanism is that widening search too much may dilute the focus and quality of applications, making matches less effective.

The essay makes two main contributions. Substantively, it provides evidence that caseworkers within public employment services influence one another's behavior, with measurable consequences for job seekers. Conceptually, it shows that policy delivery is socially embedded: the effects of job-search assistance depend not only on formal rules or program resources, but also on workplace norms inside the institutions that implement it. More broadly, these findings show that variation in frontline implementation is central to understanding the effectiveness of public employment services.

Essay 2: Gender norms in referrals and labor market outcomes

Beyond peer environments, caseworker behavior may also reflect broader social and cultural norms, with important consequences for labor market inequality. Occupational segregation remains a persistent feature of modern labor markets and a major driver of gender inequality, despite earlier gains in female labor force participation and occupational integration (Blau et al., 2013; Blau and Kahn, 2017; Levanon et al., 2009). A growing literature argues that this persistence reflects not only preferences and comparative advantage, but also gender norms and institutional frictions that shape occupational sorting through both worker and employer behavior (Sorkin, 2017; Vattuone, 2024; Cejka and Eagly, 1999; Barbulescu and Bidwell, 2013; Azmat and Petrongolo, 2014; Baert, 2018; Kuhn and Shen, 2023; Card et al., 2025). Building on this insight, the second essay, ***"Returns to Breaking the Norm: Evidence from Job Referrals,"***, examines whether public policy can mitigate occupational segregation through guidance during job search.

Using rich Swedish administrative data covering all unemployment spells from 2007–2017, the essay studies whether exposure to less gender-stereotypical caseworkers affects job seekers' occupational choices and labor market outcomes. Identification comes from quasi-random assignment of job seekers to caseworkers based on date-of-birth rules within the Public Employment Service, following the empirical approach of Cederlöf et al. (2025). The analysis shows that assignment to less gender-stereotypical caseworkers leads to faster exits from unemployment and higher post-employment earnings, with effects that persist in the medium and long run. The effects are also gender-asymmetric, with men benefiting more from being assigned to a more non-stereotypical caseworker.

The essay offers two main contributions. On the one hand, it identifies a policy-relevant mechanism through which occupational segregation can be reduced at the job-search stage. On the other hand, it complements existing work on labor market sorting (Sorkin, 2017; Card et al., 2025) by showing that labor market intermediaries, not only worker preferences or employer

behavior, help shape economic outcomes. In this sense, the essay highlights that social norms are economically consequential and, when they are embedded in institutional guidance, they become part of policy delivery. The findings further suggest that even after accounting for other productive caseworker traits or strategies, non-stereotypical guidance exerts an independent effect, underscoring that the effectiveness and equity implications of labor market policy depend in part on the norms and practices of frontline staff.

Essay 3: Public financing of secondary education in Bangladesh

The final essay turns from labor market policy to education finance in a developing-country context. It asks why substantial public support may fail to improve core outcomes when implementation is weak. Titled ***"Financing Secondary Education: Closing Gaps or Boosting Grades?"***, the essay examines whether increased public funding translates into better student performance. This question remains central in education economics, where a large literature shows that the returns to educational spending depend critically on institutional quality, governance, and accountability (Hanushek, 2006; Wößmann, 2003; Duflo et al., 2015; Lafortune et al., 2018).

The essay evaluates a large-scale payroll subsidy program, called Monthly Pay Order (MPO), for secondary schools in Bangladesh using a regression discontinuity design based on government-set eligibility thresholds. By comparing schools just above and below the cutoffs, the analysis identifies local average treatment effects of the subsidy on school-level outcomes, with standard validity tests supporting the design following Cattaneo et al. (2020). The empirical analysis combines a novel administrative dataset with regional indicators of poverty and corruption, as well as event-level information from GDELT (Leetaru and Schrodt, 2013), to explore potential mechanisms.

The results show no significant effects of the subsidy on student academic performance, measured by passing rates and the share of distinctions in the Secondary School Certificate examination. There is only weak evidence of a positive effect on girls' secondary school participation. These findings suggest that while additional funding may modestly expand access for girls, it does not necessarily improve learning outcomes. Consistent with the broader literature, the limited impact appears to reflect weak incentives and accountability mechanisms (Wößmann, 2003; Duflo et al., 2015; Lafortune et al., 2018). The results also show little heterogeneity across poorer versus richer regions or across areas with differing levels of corruption.

To investigate mechanisms, the essay examines teacher staffing, class size, and complementary evidence on teacher dissatisfaction. Mediation analysis shows no meaningful changes in staffing or class sizes, suggesting limited improvements in instructional capacity. Evidence from an alternative dataset further points to rising dissatisfaction among MPO-supported teachers. Al-

though the available data do not allow a direct assessment of how resources were reallocated or how teacher dissatisfaction affected performance, these patterns are consistent with the null effects. More broadly, the findings suggest that increased public spending is unlikely to improve outcomes unless implementation structures, accountability, and incentives translate funding into meaningful changes in school functioning.

This essay offers an important counterpoint to the labor market chapters. It demonstrates that increased public spending on services or institutions does not automatically lead to better outcomes, especially when implementation is weak or when additional resources fail to influence the behavior of the key actors involved. Policymakers therefore need to look beyond inputs and consider organizational incentives, accountability, and the social processes of delivery.

Connecting the essays: a common framework

Although the three essays are situated in different empirical domains, they are connected by a common framework. Each studies a policy that is well motivated in principle. Job-search assistance is widely considered an effective way to reduce unemployment. Guiding job seekers toward broader opportunities could plausibly improve matching. Reducing gendered barriers to occupational mobility could improve both equity and efficiency. Financing teacher payrolls could plausibly strengthen schools and improve education. Yet in each case, the dissertation asks whether these intuitive expectations hold once the policy is viewed through the lens of implementation and institutional context.

A useful way to think about the dissertation is therefore to distinguish between *policy intent*, *policy delivery*, and *policy effect*. Policy intent refers to the formal objective: reducing unemployment, improving matching, broadening opportunity, or strengthening schooling. Policy delivery refers to the processes through which the policy reaches individuals: caseworker referrals, counseling, assignment rules, subsidy disbursement, and the organizational settings in which these occur. Policy effect refers to the observed outcome. The central claim running through the dissertation is that the mapping from intent to effect is mediated by delivery. Peer norms, gender stereotypes, discretion, and institutional quality all shape that mapping.

This framework also helps explain why the three essays belong together. They are not separate policy evaluations joined only by method. Rather, they are connected investigations of how institutions shape behavior and how that, in turn, shapes economic outcomes. The dissertation is therefore concerned as much with the mechanisms of implementation as with the causal effects themselves.

Concluding remarks

The overarching contribution of this dissertation is to show that policy outcomes are shaped by more than program design and financial resources. They are mediated by caseworkers, peers, gender norms, administrative rules, and institutional quality. In the labor market context, these mediating factors influence how unemployed individuals are guided toward jobs and occupations. In the education context, they shape whether public spending translates into improved learning or merely sustains existing structures.

By bringing these cases together, the dissertation argues for a broader approach to policy evaluation. Effective policy analysis should ask not only whether an intervention exists and how much it costs, but also how it is delivered, by whom, and under what conditions. Policies do not act on empty space. They operate through institutions, and institutions are populated by individuals whose behavior is shaped by social and organizational forces. Understanding those forces is essential for explaining why some policies succeed, why others fail, and how public interventions can be designed to improve economic outcomes more effectively.

References

Azmat, G. and Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour economics*, 30:32–40.

Baert, S. (2018). Hiring a gay man, taking a risk?: A lab experiment on employment discrimination and risk aversion. *Journal of homosexuality*, 65(8):1015–1031.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, 77(4):1047–1094.

Barbulescu, R. and Bidwell, M. (2013). Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science*, 24(3):737–756.

Belot, M., Kircher, P., and Muller, P. (2019). Providing advice to jobseekers at low cost: An experimental study on online advice. *The review of economic studies*, 86(4):1411–1447.

Belot, M., Kircher, P., and Muller, P. (2025). Do the long-term unemployed benefit from automated occupational advice during online job search? *The Economic Journal*, page ueaf041.

Beugnot, J. and Peterlé, E. (2020). Gender bias in job referrals: An experimental test. *Journal of Economic Psychology*, 76:102209.

Blau, F. D., Brummund, P., and Liu, A. Y.-H. (2013). Trends in occupational segregation by gender 1970–2009: Adjusting for the impact of changes in the occupational coding system. *Demography*, 50(2):471–492.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, 55(3):789–865.

Card, D., Colella, F., and Lalive, R. (2025). Gender preferences in job vacancies and workplace gender diversity. *Review of Economic Studies*, 92(4):2437–2471.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):F452–F477.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cederlöf, J., Söderström, M., and Vikström, J. (2025). The role of caseworkers: Job finding, job quality and determinants of value-added. *Journal of the European Economic Association*, page jvaf055.

Cejka, M. A. and Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and social psychology bulletin*, 25(4):413–423.

Cheung, M., Egebark, J., Forslund, A., Laun, L., Rödin, M., and Vikström, J. (2024). Does job search assistance reduce unemployment? experimental evidence on displacement effects and mechanisms. Technical report, Working Paper.

Crépon, B., Dejemeppe, M., and Gurgand, M. (2005). Counseling the unemployed: does it lower unemployment duration and recurrence? *IZA discussion paper*.

Del Carpio, L. and Guadalupe, M. (2022). More women in tech? evidence from a field experiment addressing social identity. *Management Science*, 68(5):3196–3218.

Dhia, A. B., Crépon, B., Mbih, E., Paul-Delvaux, L., Picard, B., and Pons, V. (2022). Can a website bring unemployment down? experimental evidence from france. Technical report, National Bureau of Economic Research.

Duflo, E., Dupas, P., and Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from kenyan primary schools.

Journal of public Economics, 123:92–110.

Dustmann, C., Rajah, N., and Van Soest, A. (2003). Class size, education, and wages. *The Economic Journal*, 113(485):F99–F120.

Hanushek, E. A. (2006). School resources. *Handbook of the Economics of Education*, 2:865–908.

Hederos, K., Sandberg, A., Kvissberg, L., and Polano, E. (2025). Gender homophily in job referrals: Evidence from a field study among university students. *Labour Economics*, 92:102662.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The quarterly journal of economics*, 115(4):1239–1285.

Kuhn, P. and Shen, K. (2023). What happens when employers can no longer discriminate in job ads? *American Economic Review*, 113(4):1013–1048.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, 10(2):1–26.

Leetaru, K. and Schrodt, P. A. (2013). Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, volume 2, pages 1–49. Citeseer.

Levanon, A., England, P., and Allison, P. (2009). Occupational feminization and pay: Assessing causal dynamics using 1950–2000 us census data. *Social forces*, 88(2):865–891.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives*, 14(3):115–136.

Marshall, J. H. (2009). School quality and learning gains in rural guatemala. *Economics of Education Review*, 28(2):207–216.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–145.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World development*, 22(9):1325–1343.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, 116(2):681–704.

Sorkin, I. (2017). The role of firms in gender earnings inequality: Evidence from the united states. *American Economic Review*, 107(5):384–387.

Van den Berg, G. J., Kjærsgaard, L., and Rosholm, M. (2014). To meet or not to meet, that is the question: short-run effects of high-frequency meetings with case workers. Technical report, Working Paper.

Van der Klaauw, B. and Vethaak, H. (2022). Empirical evaluation of broader job search requirements for unemployed workers. Technical report, JSTOR.

Vattuone, G. (2024). Worker sorting and the gender wage gap. Working paper.

Weber, A. and Hofer, H. (2003). Active job-search programs a promising tool? a microeconometric evaluation for austria. Technical report, Reihe Ökonomie/Economics Series.

Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford bulletin of economics and statistics*, 65(2):117–170.

Vacancy Referrals and Broader Job Search: The Role of Caseworkers and Peer Spillovers[1]

[1]I would like to thank Johan Vikström, Peter Fredriksson, Pierre Deschamps, Olof Rosenqvist, Oskar Nordström Skans, Lena Hensvik, Georg Graetz, Jonas Cederlöf, Andrei Groshkov, and Eva Mörk for their insightful comments and suggestions. This work has further benefited from discussions with members of the Uppsala Labor Group and participants at the CREST–IFAU Workshop.

Introduction

Job-search assistance (JSA) is perhaps the most efficient policy tool in reconnecting the unemployed with employment. Delivered through the caseworkers at the public employment service (PES) offices, JSA typically involves vacancy referrals[2], counseling, job suggestions, and career planning that are tailored to individual needs and shaped by caseworker discretion. There is strong empirical evidence supporting the effectiveness of JSA in increasing re-employment probabilities (Card et al., 2010; Weber and Hofer, 2003; Crépon et al., 2005; Cheung et al., 2024; Van den Berg et al., 2014). Consistent with this rationale, many developed countries expand spending on job-search assistance and public employment services during periods of rising unemployment.[3] Some recent studies have also examined the influence of caseworker attributes (Cederlöf et al., 2021), caseworker-job seeker interactions (Schiprowski, 2020), and referral mechanisms (Bollens and Cockx, 2017) that have proved to be very relevant for job seeker's labor market outcomes. However, little is known about the factors that drive individual caseworker's vacancy referral strategy or behavior and how these, in turn, affect the labor market prospects of job seekers.

Vacancy referrals constitute a core component of job-search assistance, serving as one of the most direct and effective tools through which caseworkers at public employment service (PES) offices facilitate job matching. Cheung et al. (2024) show that increased number of referrals significantly improves re-employment outcomes of the job seekers. In addition to the frequency of referrals, the type of referral may also be a key factor. Recent policies in many countries encourage job seekers to consider a wider set of occupations and opportunities beyond their previous occupations or narrow preferences, often termed as "broader referral". This strategy aims to broaden job seekers' search horizons to improve match quality and expedite re-employment, assuming that unemployed workers often misjudge their employment prospects (Van der Klaauw and Vethaak, 2022). Empirically, Belot et al. (2019; 2022) document that broader referrals improve labor market outcomes for underprivileged groups of job seekers. However, recent evidence reports mixed effects of broader referrals—ranging from muted to negative impacts on employment outcomes (Dhia et al., 2022; Van der Klaauw and Vethaak, 2022).

In this paper, I explore two under-researched questions related to JSA and vacancy referrals. Firstly, I investigate whether an individual caseworker's vacancy referral behavior or strategy is influenced by their peers at the local PES office. This inquiry is grounded in a broad literature on peer effects, which highlights how individuals are often shaped by the actions and social norms of

[2]A vacancy referral is defined as an instance in which a caseworker formally assigns or recommends a job seeker to a specific job opening with a particular employer.

[3]Appendix figure A1 shows the relationship between unemployment rates, and PES administrative and JSA expenditures using OECD country-year observations from 2005-2020. The figure indicates a positive correlation, consistent with countercyclical increases in job-search support related spending during periods of higher unemployment.

those around them (Sacerdote, 2001; Mas and Moretti, 2009; Bandiera et al., 2009). While prior studies on active labor market programs (ALMPs) have focused extensively on the effectiveness of program participation for job seekers, no attention has been paid to within-office behavioral dynamics among caseworkers, specifically, how caseworkers adapt their work strategies in response to peer norms. This study exploits the variation in practice-environment from caseworkers moving between local offices to identify whether and how caseworker referral behavior adjusts to the peer environment. To the best of my knowledge, this study provides the first causal evidence of peer-driven behavioral spillovers among caseworkers.

Secondly, this paper examines how caseworker's broader referral behavior affects their job seekers' labor market outcomes. Unlike prior research (Belot et al., 2019, 2022; Dhia et al., 2022) that evaluates broader search behavior through algorithmic job suggestions or informational nudges, typically implemented on online platforms and often targeted at narrowly defined and highly disadvantaged groups such as individuals with disabilities or long unemployment duration, this study analyzes vacancy referrals that PES caseworkers routinely issue to job seekers as part of their standard services. I identify the causal effect of shifts in the broader referral environment on job seekers' reemployment outcomes by exploiting within-caseworker variation before and after relocation within local PES offices. The findings offer new insights about the role of broader referrals on job seekers' reemployment prospects.

This study combines rich administrative data from the Swedish PES offices with matched employer–employee records, wage data, and demographic information from Statistics Sweden, covering 2004–2018. These data allow tracking of both caseworkers' referral behavior and their job seekers' labor market outcomes. The empirical strategy employs a difference-in-differences and event-study design, focusing on PES caseworkers who move between local offices. Variation in the "Leave-one-out" mean[4] of broader referral behavior across origin and destination offices, defined as exposure or treatment to the caseworkers, identifies how differences in practice environments affect the "migrant" caseworker's[5] broader referral behavior. The key identifying assumption in this setting requires that no other factors except the practice-environment are affected or changed simultaneously with the move of caseworkers that are correlated with the changes in practice-environment and affect caseworkers' referral strategy. I use a similar specification to estimate reduced-form effects comparing the labor market outcomes of a migrant caseworker's new and previous job seekers. This approach enables a causal interpretation of how changes in caseworkers' broader practice environments affect their clients' reemploy-

[4]The counterfactual office environment is the average broader referral rate among other caseworkers in the office—at the destination office before the move and at the origin office after the move.

[5]The term "migrant" caseworker refers to caseworkers who move between offices, not to those with migrant citizenship status.

ment prospects, thereby shedding light on the effectiveness of broader relative to narrower vacancy referrals. The empirical design follows Molitor (2018) and Avdic et al. (2023), who study analogous spillovers among physicians.

I find evidence of strong spillover effects. Caseworkers who move to offices with a higher rate of broader referrals significantly increase their own broader referral behavior in response to the new practice environment. The results suggest that a one percentage point (pp) increase in the office's broader referral rate leads to a 0.346 pp increase in the migrating caseworker's broader referral rate. Event-study results further validate the presence of a behavioral discontinuity following relocation. A potential concern in estimating peer effects is the reflection problem (Manski, 1993), which I address by defining the practice environment prior to relocation and control for exogenous peer characteristics to isolate genuine peer influences. Overall, the results are robust across all specifications, indicating that the estimated spillovers capture behavioral adaptation rather than correlated shocks or simultaneity. I also decompose the overall spillover effect into peer influence versus office-level or local labor market characteristics. This decomposition suggests that peer-group influence (0.284) is much stronger than office-specific elements (0.062).

In terms of job seeker's outcomes, I find that when a migrant caseworker is exposed to an office with a 10 pp higher broader referral intensity, the probability of exiting unemployment within 90 and 180 days decreases by 1.4 and 1.5 pp respectively. Event-study dynamics confirm that these discontinuities are both statistically significant and persistent. One plausible reason for this decline might be the fact that broader referrals induce job seekers to submit a larger number of applications to a wider set of vacancies, potentially lowering average application quality. If this occurs, such behavior may weaken the signaling value of individual applications to employers, reducing callback and hiring probabilities. Hence, a broader referral strategy may thus increase search scope at the expense of credibility and match precision, leading to slower reemployment transitions despite higher nominal search activity. The results of this paper provide robust evidence that behavioral shifts in caseworker referral strategies, driven by changes in the surrounding practice environment, can translate into meaningful labor market consequences. The impact on jobseekers' outcomes is robust to controlling for local labor market-by-year fixed effects; and remain stable after incorporating these controls. Using the full sample, I further compare migrant and non-migrant (stayer) caseworkers and find no significant differences between movers and stayers.

This paper relates to several strands of literature. First, this paper contributes to the literature on peer effects and behavioral spillovers across institutional contexts. In health care, Molitor (2018) and Avdic et al. (2023) show that cardiologists adjust treatment styles after moving between hospitals, influenced by local practice norms. In labor, Ichino and Maggi (2000) document peer effects on shirking, while Sacerdote (2001) finds that randomly assigned college roommates affect academic and social outcomes. Åslund and Fredriksson

(2009) further show that ethnic peers' welfare use influences individual take-up under Sweden's refugee placement policy. Building on these studies, I provide the first causal evidence of peer-driven behavioral spillovers among PES caseworkers. By exploiting plausibly exogenous variation from caseworkers' transitions across local offices in Sweden, I show that exposure to new peer environments shifts individual broader referral behavior. This demonstrates that referral behavior is adaptable rather than predetermined, and highlights the influence of peer norms in shaping the delivery of public employment services.

Secondly, using rich administrative data with detailed caseworker and job seeker information, this study contributes to the literature on broader job search policies. Belot et al. (2019) show that algorithm-based job recommendations increase interview rates, particularly among the long-term unemployed, and Belot et al. (2022) find lasting employment and earnings gains. Both rely on small-scale experiments with limited external validity. In contrast, Dhia et al. (2022) find no effects of large-scale online recommendations, while Altmann et al. (2018) report modest gains from informational brochures concentrated among high-risk job seekers. The only study, to my knowledge, evaluating mandated broader search is Van der Klaauw and Vethaak (2022), who find that enforcement tied to vacancy referrals lowers job-finding rates. However, this paper adds nuance by examining broader search arising from institutional shifts in caseworker behavior and shows that when caseworkers move to offices with higher broader-referral intensity, their job seekers' exit probabilities decline, which suggests that indiscriminate broadening can hinder reemployment.

Thirdly, this paper relates to the literature on the role of caseworkers in shaping job seeker outcomes. Prior studies offer mixed evidence. Using Swiss data, Lechner and Smith (2007) find that caseworkers do not outperform algorithmic assignment of job seekers to ALMPs, whereas Cederlöf et al. (2021) show that caseworker characteristics affect job seekers' labor market performance. Likewise, Schiprowski (2020) and Van den Berg et al. (2014) find that more frequent meetings shorten unemployment spells. The present study extends this literature by documenting behavioral spillovers in caseworkers' broader referral strategies and their implications for job seeker outcomes, offering an organizational perspective on public employment service effectiveness.

Finally, this paper contributes to research on vacancy referrals, their channels and effectiveness. Cheung et al. (2024) show that referrals and individualized meetings are the main channels through which caseworkers affect unemployment duration, while Bollens and Cockx (2017) find that personalized referrals outperform automated ones. This study shifts the focus from the quantity, or mode of referral delivery, to the type of vacancies referred.

The remainder of the paper is organized as follows. The next section provides institutional background on the Swedish Public Employment Service and caseworkers' broader referral behavior. The subsequent sections describe the data and sample selection, discuss caseworkers' practice environments, and present the empirical framework. The main results and robustness checks

are then presented, followed by an examination of job seekers' outcomes and heterogeneity. The final section concludes by summarizing the findings and discussing their implications.

Swedish Public Employment Service offices and caseworkers

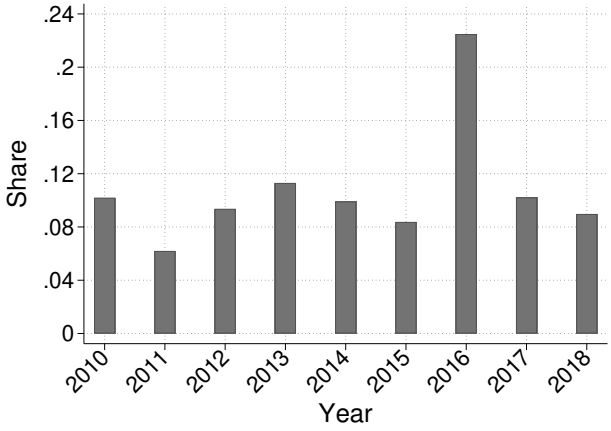
The Swedish Public Employment Service (PES, *Arbetsförmedlingen*) is a central institution in Sweden's labor market policy, tasked with reducing unemployment and supporting job seekers in their transition to work by ensuring efficient matching between job seekers and employers while addressing structural challenges such as skill mismatches, regional disparities, and the integration of under-represented groups, including migrants and individuals with disabilities. PES offers extensive services to job seekers (e.g., vacancy referral, career counseling, training programs, and financial support for skill development) and to employers, who benefit from recruitment assistance, hiring subsidies, and access to a large candidate database; notably, more than half of PES staff are caseworkers. Although guidelines and legislation frame service provision, caseworkers retain substantial discretion, a flexibility attributed to broad policy frameworks and outcome-focused performance evaluations rather than mandated strategies (Lundin, 2004; Lagerström, 2011). This autonomy allows them to tailor interventions to local labor market conditions and heterogeneous client needs. Sweden also imposes no standardized educational requirements for caseworkers, instead valuing diverse professional backgrounds, and local PES offices are encouraged to adapt their organizational practices to regional circumstances—for example, matching job seekers to caseworkers based on sector expertise or using neutral assignment rules such as date of birth to balance caseloads (Lundin and Thelander, 2012). These varied approaches underscore the foundational importance of caseworkers in implementing labor market policy and highlight the need to understand and strengthen the factors that enhance their effectiveness.

Caseworker movement within Swedish PES

Caseworker mobility within public employment systems presents a unique opportunity to examine how individual behaviors adapt to changing organizational environments. This geographic relocation of caseworkers across local offices can generate quasi-experimental variation in exposure to peer environment. Importantly, the localized nature of PES offices, each with potentially distinct referral tendencies and managerial styles, makes caseworker relocation a valuable source of identifying variation.

To identify caseworkers who relocate their job, I focus on those who move across two different offices. Moves are identified using office and caseworker identifiers. I treat caseworkers as movers who have worked at a single office

Figure 1. Annual transition rate of caseworkers across offices



Notes: Figure 1 is drawn using the PES admin dataset. This figure shows the annual transition rate among caseworkers in Sweden for the period 2010-2018. This figure represents those caseworkers, who held at least 1 physical meeting with their job seekers at the office.

in a specific year and who have not switched offices more than once during the entire sample period. Restricting the analysis to individuals with a single move sharpens identification of peer spillover effects. Multiple relocations expose individuals to varying peer environments over short periods, obscuring the influence of any specific practice environment. Repeated movers may also differ in unobserved traits (e.g., adaptability or network formation) that affect both mobility and responsiveness to peer norms. Focusing on one-time movers therefore enables cleaner attribution of spillover effects to a single peer environment rather than to selection-driven mobility patterns. Of the 11,352 unique caseworkers analyzed, 1,789 (15.76%) were identified as movers.

Figure 1 highlights the migration patterns of caseworkers working in Swedish public employment service offices during 2010-2018. It shows that the annual transition rate of caseworkers in Sweden. For most part of the last decade, the transition rate remains in the range of 8%-10%. However, there is a spike around 2016 when the transition of caseworker peaked to more than twenty percent. Most of the movement in this transition rate could be attributed to the gradual downsizing of local PES offices in Sweden during the sample period.

Broader referral behavior

In this study, broader referral behavior is defined by the extent to which the initial job suggestion or vacancy referral made by a caseworker falls outside the occupational preference informed by the job seeker during their registration at the PES. These preferences could be shaped by a combination of factors, including the job seeker's ideal or "dream" job, perceived employment prospects,

and alignment with prior experience or qualifications[6]. This measure captures instances where the caseworker recommends positions that do not align closely with the job seeker's stated preferences. I track this broader vacancy referral behavior of individual caseworker over time and across different offices using the unique caseworker identification number. This unique identifier remains consistent throughout a caseworker's tenure during the sample period. I also focus on the first referral, rather than alternative measures such as the average rate of broader referrals or probability of ever receiving a broader referral. While many factors might influence the referral behavior of caseworkers during a job seeker's unemployment spell, focusing on the first referral helps to some extent in reducing concerns about endogeneity in the vacancy referral behavior of the caseworkers. Throughout this paper, the terms referral and broader referral are used interchangeably.

Understanding broader referral behavior is important from both a behavioral and policy perspective. These referrals can influence the quality of job matches, job seeker motivation, and ultimately labor market outcomes. While previous studies have shown that informational nudges or voluntary broadening of search can improve job outcomes (Belot et al., 2019, 2022) for selective job seekers but not for others (Dhia et al., 2022), mandatory broadening might be counterproductive for the job seekers (Van der Klaauw and Vethaak, 2022). Overall, the evidence on the effectiveness of broader vacancy referral remains limited. This paper provides new evidence on how caseworkers' adoption of broader referral behavior—triggered by changes in their peer environment—affects job seekers' employment outcomes in a context where vacancy referral is institutionalized. Here, the analysis contrasts the effect of highly broader practice-environment with less broad ones, rather than assessing the impact of broader referrals on their own.

Data and sample

For this study, I use different Swedish administrative datasets containing important demographic, education and employment history of the unemployed job seekers registered at the PES. I use the data from 2004-2018 as there is not much referral data prior to 2004; whereas, I do not have information beyond 2018. I use the population register Louise from Statistics Sweden that contains demographic information (e.g., age, gender etc.) and education (type and level). I also use the matched employer-employee data (RAMS) from Statistics Sweden to add information of earnings and employment status of the job seekers

[6]Appendix figure A2 examines whether job seekers' stated preferences, at the PES, genuinely reflect their actual occupational preferences. To assess this, I predict the probability that each job seeker prefers a given occupation and then rank these predicted preferences. Approximately 60% of job seekers' stated preferences fall within their top four predicted occupations, suggesting that the stated preferences largely capture their true occupational choices.

prior and after each unemployment spell. In this way, I am able to extract the rich employment history of each job seeker.

PES register data is used to construct unemployment spell data. The same register is also used to construct the vacancy referral dataset, which contains the actions of both job seekers and their assigned caseworkers for each vacancy referral and job suggestion. This dataset contains the vacancy identifier along with location of the vacancy, occupational code, type of vacancy, application deadlines, information about whether the job seeker applied to or not etc. To elicit the preference[7] of the job seekers as well as referral behavior[8] of the caseworkers, I use the information from the vacancy referral dataset. I also use meeting information from the PES registers at the PES to extract the caseloads (e.g., number of job seekers assigned to the caseworkers), frequency of vacancy referral and job suggestions per year, tenure of caseworkers (i.e., the number of years worked as caseworkers at the PES), the number of offices at which a caseworker worked at in a given year and most importantly, the point in time when a migrant caseworker switches office. The major outcome variables used in this study are: (i) probability of exiting unemployment from receiving the first referral within 90 and 180 days respectively, and (ii) log earnings in year $t+1$ and $t+2$ after the spell starts, or specifically the short-run future earnings. To map the local labor market within my sample dataset, I use the local labor market definition[9] from Statistics Sweden.

The sample consists of all job seekers registered at the PES from 2004-2018. As stated earlier, I exclude caseworkers who switched offices more than once during their tenure. To enhance the precision of the empirical analysis, the sample is restricted to PES offices that employ at least three caseworkers within a given year. This restriction is motivated by the need to reduce measurement error in the construction of the leave-one-out mean of referral intensity, which serves as a proxy for average peer behavior within an office. All movers in the sample have worked at least 1 year in origin and destination offices respectively. The final dataset contains 626,248 unique unemployment spells (splitted between stayers and migrant caseworkers), 1,789 unique migrant caseworkers, and 9,563 unique stayers.

[7] Job seekers are allowed to state their four digit occupational preferences after the of registration at PES.

[8] By referral behavior, I mean both job suggestion and job referral towards a job seeker.

[9] In my sample, I fix the local labor market using the 2005 definition from Statistics Sweden. The local labor market mappings from Statistics Sweden are compiled on yearly basis; however, the mappings do not vary much across years.

Practice-environment based on broader referral

Practice-environment

Building on the definition of broader referral behavior in the earlier section, the empirical analysis seeks to identify how migrant caseworkers adjust their own referral behavior in response to changes in their surrounding practice environment. I calculate the broader referral behavior for each caseworker at office-year level using a "leave-one-out" method that excludes their own job seekers. While the practice environment likely includes unobserved components, the leave-one-out measure provides a tractable proxy by eliminating the mechanical correlation between an individual's behavior and the office-level average.[10] This strategy allows for a cleaner estimation of the extent to which caseworkers internalize and respond to the referral norms of their new organizational setting. I define practice-environment of caseworker $n \in N$ in office $z \in Z$, where job seeker $m \in M_{zt}$ received a broader referral (as their first referral) in time period $t \in T$, as the ratio:

$$\Xi_{nzt} = \frac{\sum_{m \in M_{kzt}} \mathbb{1}(Broader_m = 1)}{M_{kzt}}, \qquad \forall k \neq n \in N \tag{1}$$

where $M_{kzt} \subset M_{zt}$ is the subset comprising all job seekers other than those assigned to caseworker n. Hence, Ξ_{nzt} is caseworker n's exposure to the practice-environment with respect to the rate of sending out broader referral to the relevant job seekers in office z at time t. Next, I define the difference in these practice-environments between a migrating caseworker's origin ($z_{Ori.n}$) and destination ($z_{Des.n}$) offices at a given point in time as:

$$\Delta_{nt} = \Xi_{z_{Des.n}t} - \Xi_{z_{Ori.n}t} \tag{2}$$

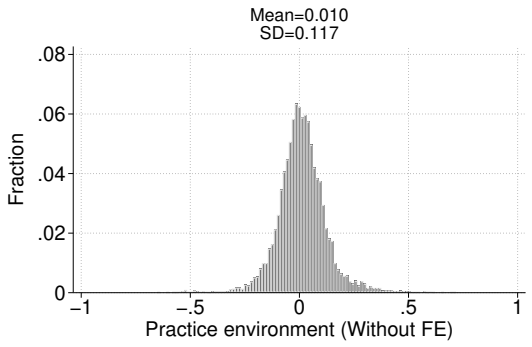
In essence, Δ_{nt} is the period-specific difference in "leave-one-out" shares between the offices where caseworker n worked before and after relocation. At each point in time, one of the two shares in equation (2) is counterfactual, defined as the peer exposure generated by caseworkers in the office where n is not currently employed. Together, equations (1) and (2) define caseworkers' exposure to the practice environment over time. The difference in equation (2) provides the time-varying measure of this exposure.

This rich administrative data sample generates substantial variation in exposure to different practice-environments, which are measured using the leave-one-out share of broader referrals at the office-year level.[11] To illustrate how

[10] I explore office-level practice-environment instead of regional variation in practice-environment for two important reasons: (i) the sample dataset can not identify the location of office; (ii) office-level practice environments enables more precise identification, as local office cultures and norms may vary meaningfully even within the same geographic area.

[11] Appendix figure A3 shows the distribution of practice-environments for all the migrant caseworkers in the sample dataset, which shows the variation in levels.

Figure 2. Time-Variant Change (Δ_{nt}) in Practice Environment for Migrating Caseworkers



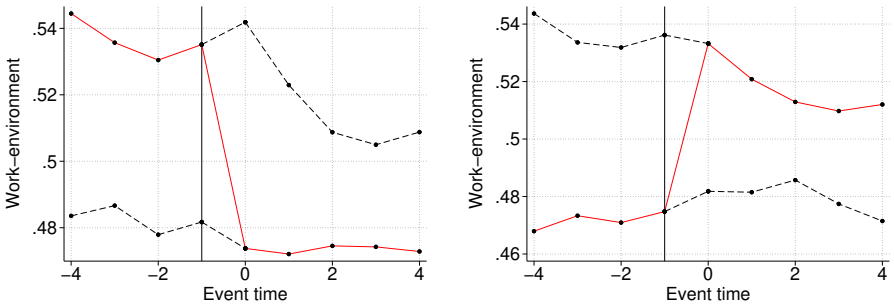
Notes: Figure 2 shows the distribution of changes in practice environment for migrant caseworkers due to inter-office moves during 2004-2018, and it uses year-specific referral environments at the time of the move. Measures are constructed from leave-one-out broader referral rates.

these environments change when caseworkers relocate, figure 2 shows the distribution of changes in broader referral intensity (Δ_{nt}). The distribution is centered around zero, indicating that movers are about equally likely to experience increases or decreases in referral intensity. Overall, the figure demonstrates ample variation in the changes in caseworker's practice environments, which is crucial for identifying the causal effects of exposure to different levels of office norms on caseworker referral behavior and job seeker outcomes.

Figure 3, on the other hand, depicts the evolution of referral intensity, measured as the share of broader referrals, at different types of PES offices before and after a caseworker relocates. The analysis distinguishes moves from high- to low-referral offices (Panel 3a) and from low- to high-referral offices (Panel 3b). For each group, the solid lines plot the realized practice environment (the leave-one-out referral intensity at the office where the caseworker is located), while the dashed lines show the corresponding counterfactual trend in the origin or destination office.

Both panels reveal a clear pattern: despite some year-to-year variation, offices maintain their relative ranking in referral intensity. High-referral offices remain high-referral after receiving new staff, and low-referral offices remain low-referral, indicating that referral culture is a persistent institutional characteristic rather than one that shifts quickly with staff mobility. Given the temporal fluctuations in office styles, it is important to account for deviations in referral intensity (Δ_{nt}) within the empirical strategy. Including Δ_{nt} as a regressor allows the model to absorb linear divergence in office-level practices that might otherwise confound treatment effects, ensuring that observed behavioral changes are not spuriously attributed to independent shifts in local norms. Overall, figure 3 demonstrates that office environments evolve but remain relatively stable in their relative intensity. This persistence motivates focusing on changes in

Figure 3. Dynamics of Referral Intensity in Different Office Types
(a) Move from High- to Low-Referral Office *(b)* Move from Low- to High-Referral Office



Notes: Figure 3 is based on sample data from 2004–2018. Solid lines show the realized office-level referral intensity; dashed lines show the counterfactual (origin or destination) intensity. Event time is represented in the horizontal axis and is centered on the relocation year.

caseworker exposure to distinct referral cultures and incorporating time-varying adjustments to isolate their causal effects.

Econometric framework and identification

The empirical strategy in this paper aims to quantify the impact of caseworkers' practice-environment on their individual decision making styles in the context of sending out broader vacancy referral to their respective job seekers. The approach exploits plausibly exogenous variation arising from the mobility of migrant caseworkers across offices and relies on a difference-in-differences and event-study framework. The identification leverages variation in the practice environment across origin and destination offices to estimate the causal effect of changes in caseworkers' local environments on their own treatment behavior following a move. The major identifying assumption is that no other factors except the practice-environment are affected or changed simultaneously with the move of caseworkers that are correlated with the changes in practice-environment and affects caseworkers' treatment styles. A similar econometric framework is also employed to examine the effect of the practice environment on job seekers' subsequent labor market outcomes.

Spillover effects among migrant caseworkers

The empirical specification in this paper is based on the unemployment spell-level difference-in-differences (DiD) framework for job seeker $m \in M$, assigned to caseworker $n \in N$ at time $t \in T$ that is given in the following equation:

$$y_{mnt} = \beta \Delta_{nt} + \phi(\Delta_{nt} \times \text{Post}_t) + \Psi'_{mnt}\Omega + \zeta_n + \zeta_t + \varepsilon_{mnt} \tag{3}$$

where the outcome y_{mnt} is a binary variable that takes 1 if a job seeker receives a broader referral and 0 otherwise. The dummy, $Post_t$, is an indicator variable that assumes 1 for all subsequent time periods after the caseworker n's migration to a new office, i.e., $Post_t = \mathbb{1}_{t \geq t_0}$ where t_0 is the time when the migration takes place. This baseline model also includes caseworker and referral year fixed effects (ζ_n and ζ_t) along with controls for job seeker characteristics, Ψ_{mnt}. The characteristic controls and fixed effects are used to adjust for the observed and unobserved heterogeneity across job seekers, caseworkers and time. Δ_{nt} is a continuous variable in the range [-1,1] that captures the time-variant empirical variations in the practice-environment of the caseworkers (with respect to the average share of sending out broader referrals to the job seekers registered at those offices at time t) after their relocation. The major parameter of interest is the ϕ in the baseline DiD[12] equation (3), which captures the average change in a caseworker's broader referral behavior after relocation relative to before, driven by differences in practice environments between origin and destination offices. In effect, ϕ measures the percentage-point change in a migrating caseworker's broader referral behavior associated with a one–percentage-point shift in the practice environment. I also include Δ_{nt} as a regressor to control for linear deviations in the relative trend between origin and destination practice-environments over time (see figure 3). In other words, it as a control for linear deviations in the common trend as in a standard DiD analysis.

To study the dynamics of the migrant caseworkers' behavioral responses due to the difference in the practice-environment, I expand the baseline framework into an event-study specification where the $Post_t$ dummy is replaced with period-specific indicator. as follows:

$$y_{mnt} = \beta \Delta_{nt} + \sum_{s=-T'}^{T'} \mathbb{1}(s = t') \left[\alpha_{t'} + \phi_{t'} \Delta_{nt'} \right] + \Psi'_{mnt} \Omega + \zeta_n + \zeta_t + \varepsilon_{mnt} \quad (4)$$

This dynamic pattern would help to verify the parallel trend assumption. The period-specific indicator in equations (4) is the "event-time" t', where t' is the time-index relative to the migration for caseworker n. The above equation is expected to capture the migrant caseworker's behavior in response to the changes in practice-environment Δ_{nt} over time. This is estimated by interacting the change in practice-environment (Δ_{nt}) with event time dummies $\mathbb{1}(s = t')$.

In addition to the baseline specification, I include *local labor market*[13] fixed effects to account for regional heterogeneity in job opportunities. These controls absorb structural differences across areas that may influence caseworkers' referral decisions. To further strengthen identification, I add local labor

[12] Similarly, the major parameter of interest in the even-study specification is $\phi_{t'}$.

[13] According to Statistics Sweden (SCB), a *local labor market* is a functional area centered on one or more municipalities that are largely self-sufficient in employment. A municipality qualifies as an independent local center if at least 80 percent of its employed residents work within the municipality and no external commuting flow exceeds 7.5 percent (Hedin and Tegsjö, 2006).

market–year fixed effects, which capture time-varying regional shocks (e.g., changes in unemployment or industry composition) that could otherwise confound estimated effects. Across all specifications, standard errors are clustered at the caseworker level to address serial correlation and unobserved individual heterogeneity.

Robustness analysis

One of the major challenges in estimating peer or spillover effects is the reflection problem (Manski, 1993), which is the empirical difficulty in distinguishing whether observed similarities in group behavior are due to individuals influencing each other (endogenous peer effects), due to common exposure to peer characteristics (exogenous effects), or driven by shared environments or selection into similar groups (correlated effects). However, the presence of simultaneity, where individuals both influence and are influenced by their peers, creates mechanical co-movement in outcomes that cannot be disentangled without further assumptions or structure. This limitation is especially pertinent in workplace settings, where employees are jointly exposed to organizational norms, supervision, and client characteristics. To resolve the reflection problem, I use the following identification strategies to alleviate the risk of conflating causality with correlation.

In the first specification, I use pre-move differences between origin and destination offices as a time-invariant measure. This precludes the simultaneity bias that arises when practice-environments are measured after the mover has joined the destination office. By anchoring the treatment variable to office practices before any possible influence by the mover, this strategy mimics the logic of an instrumental variable approach, using variation plausibly exogenous to the mover's own behavior. A second strategy evaluates whether spillovers are attributable to observable exogenous characteristics of the new peer group rather than their behavior. In this case, I include differences in average tenure of peers at the destination and origin offices. If more experienced peers influence behavior differently, differences in tenure could explain part of the spillover effect.

I also compare the referral behavior of migrant and non-migrant (stayer) caseworkers, following Molitor (2018). This specification includes origin and destination office fixed effects along with year fixed effects and is estimated on the full sample of movers and stayers.[14] This approach offers two advantages: (i) first, it expands the estimation sample, increasing statistical power, and (ii) second, it enables a direct comparison between movers and stayers to test for systematic differences in their respective referral intensities.[15]

[14] For stayers, the origin and destination office are defined as their current office. Their event time and change in practice environment (Δ_{nt}) are set to –1 and 0, respectively.

[15] This specification does not capture the referral behavior of individual caseworkers, as it reflects only the average behavior across groups defined by origin–destination office pairs.

Decomposition of practice-environment effect

As pointed earlier, the baseline DiD specification in this study captures the change in individual caseworker behavior due to changes in overall practice-environment after the relocation. This practice-environment could be a combined effect of the peers with whom the caseworkers work, and the office-specific elements that may also include local labor market factors. Assuming the office components are constant over time, we can disentangle the combined practice-environment effect into peer-group and office-specific effects by an inclusion of office fixed effects into the model. Including an office fixed effects in the baseline model will help to purge the effects of any unobserved difference in office-specific factors. As a result, the remaining effect in this final specification could be attributed as the average peer-group specific effect for a migrant caseworker after the relocation. The difference between the peer-group specific effects and the total effect could be interpreted as office-specific effect.

Estimating job seeker's outcomes

I use similar specifications to estimate job seekers' outcomes when their caseworkers are exposed to different offices with regard to broader referral intensity, and the reduced form specification is below:

$$y_{mnt} = \beta\Delta_{nt} + \phi(\Delta_{nt} \times \text{Post}_t) + \Psi'_{mnt}\Omega + \zeta_n + \zeta_t + \zeta_{llm\times t} + \varepsilon_{mnt} \qquad (5)$$

where y_{mnt} is the labor market outcomes of job seeker m assigned to caseworker n at time t. In this paper, two broad types of job seeker outcomes have been explored; one is probability of exit from unemployment spell (within 90 and 180 days) and the other is future earnings up to period $t+2$. Δ_{nt} is the time-variant changes in the migrant caseworker's practice-environment. The $Post_t$ is an indicator variable that assumes 1 for all subsequent time periods after the caseworker n's migration to a new office, i.e., $Post_t = \mathbb{1}_{t\geq t_0}$ where t_0 is the time when the migration takes place. Similar to the spillover estimation specification in equation 3, this specification also includes caseworker and referral year fixed effects (ζ_n and ζ_t) along with controls for job seeker characteristics Ψ_{mnt}. The job seeker's outcome model also includes local labor market by year fixed effects ($\zeta_{llm\times t}$) to account for potential contemporaneous shifts in local economic conditions. The characteristic controls and fixed effects are used to adjust for the observed and unobserved heterogeneity across job seekers, caseworkers and time.

To assess whether job seekers are systematically sorted across caseworkers, I conduct a placebo test by regressing predicted outcomes using a similar specification as in equation 5. Here, we do not control for the job seeker characteristics. Because these predicted outcomes are constructed from pre-determined characteristics, they should remain unaffected by contemporaneous referral practices for the validity of the identification. Therefore, finding

no systematic relationship between the placebo outcomes and variation in practice environments would support the assumption that job seekers are not non-randomly sorted across caseworkers.

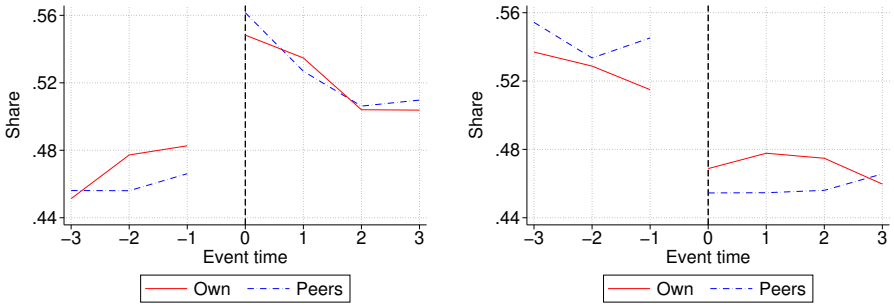
Do caseworkers change behavior upon changing offices

Figure 4 illustrates the raw association between a migrant caseworker's average referral behavior and that of their peers.[16] Figure 4a displays how a migrant caseworker's referral behavior adjusts when moving from a low-intensity to a high-intensity office in terms of sending broader referrals. Conversely, figure 4b illustrates the adjustment when transitioning from a high- to a low-intensity office. Both panels indicate that caseworkers' referral behaviors shift following relocation and tend to align with the average behavioral norm of the destination office.

Figure 4. Co-movement of Individual Work-Style with Peers

(a) Movement from Low to High Intensity Offices

(b) Movement from High to Low Intensity Offices



Notes: This figure is based on sample data from 2004–2018. Panel 4a shows how an individual caseworker's referral behavior co-moves with that of their peers when transitioning from low- to high-intensity offices. Panel 4b presents the same co-movement pattern for caseworkers relocating from high- to low-intensity offices.

Table 1 shows regression estimates of spillover effects using equation 3. Column (1) reports the baseline specification, which includes only year and caseworker fixed effects. The results indicate that a 1 pp increase in the broader referral intensity of an office leads to a 0.346 pp increase in the individual caseworker's broader referral rate following relocation. This specification

[16] Peers before event time 0 belong to the origin office, while those from event time 1 onward belong to the destination office. "Own share" refers to the average broader referral rate of the migrant caseworker, whereas "peer share" is calculated as the leave-one-out mean of the broader referral rate for each caseworker.

exploits within-caseworker variation over time, comparing referral behavior before and after a move.

Table 1. *DiD regression estimates using time-variant* Δ_{nt}

	(1)	(2)	(3)
Spillover coefficient (ϕ)	0.346***	0.291***	0.261***
	(0.058)	(0.059)	(0.062)
Covariates	✓	✓	✓
Year FE	✓	✓	✓
Caseworker FE	✓	✓	✓
Local Labor Market FE	×	✓	×
Local Labor Market × Year FE	×	×	✓
Observations	115,785	114,483	114,385
R-squared	0.056	0.059	0.068

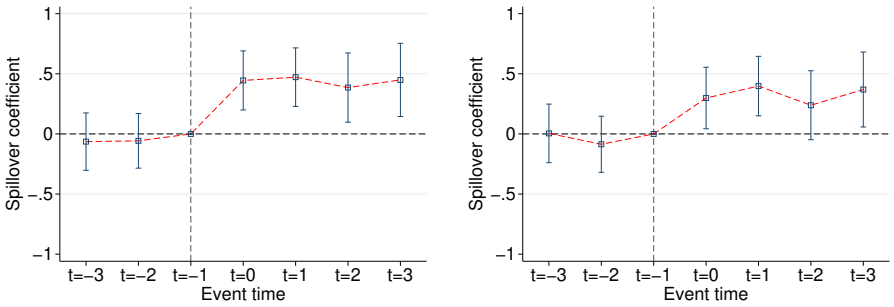
Notes: This table presents difference-in-differences estimates of the change in a caseworker's broader referral behavior across a move as a function of the time-variant change in the practice environment based on sample data from 2004-2018. Each column presents results from a separate regression. Columns are estimated over spell-level observations limited to migrant caseworkers only. The change in practice environment is defined at the office-year level. Each regression includes the same set of covariates (job seeker characteristics) and varying sets of fixed effects as shown. Standard errors are clustered at the caseworker level and shown in parentheses. Asterisks ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

However, one potential concern with the baseline specification is that unobserved spatial characteristics of the destination office, such as local economic conditions or institutional norms, may independently influence referral behavior. To address this, column (2) adds local labor market fixed effects, and the estimated spillover effect remains sizable at 0.291. Column (3) further allows local labor markets to evolve by including local labor market–year fixed effects, capturing time-varying regional shocks. The coefficient remains substantial, with a 1 pp change in the broader environment associated with a 0.261 pp change in individual behavior. Collectively, the findings indicate that caseworkers adjust their work styles to align with the prevailing practices of their colleagues and, potentially, with managerial expectations when exposed to a new practice environment. Such adaptation, consistent with theories of peer influence and social conformity, implies that professional behavior is shaped not solely by fixed individual preferences or autonomous judgment, but also by locally embedded normative pressures within organizational contexts. Moreover, a larger spillover coefficient also indicates that newly arrived caseworkers' behavior are more consistent with the prevailing norms of their destination office.

The resulting DiD estimate from table 1 are also supported by the discontinuity observed in the post-relocation period of the event-study graph in each panel of figure 5. This event-study figure compares two specifications, with each dot representing an estimated period-specific spillover coefficient $\phi_{t'}$ and vertical bars showing 90% confidence intervals. The dashed vertical line marks the event year in which caseworkers move to a new office. The dataset is

Figure 5. Dynamics of caseworker's referral behavior using time-variant Δ_{nt}

(a) CW and Year FEs *(b)* CW, Year, & LLM*Year FEs



Notes: This figure is based on the sample data for years 2004–2018. This figure shows the event-study graph of individual caseworker's behavioral response due to changes in practice-environment. Panel 5a shows the response with time-variant changes in practice-environment using caseworker and year fixed effects. Standard errors are clustered at the caseworker level. Panel 5b additionally controls for interacted local labor market and year fixed effects.

unbalanced in terms of caseworkers, and ϕ_{-1} is normalized to zero as coefficients are identified relative to this reference period. The estimates show no meaningful pre-trends: coefficients in the years preceding relocation do not differ from zero, indicating no anticipatory adjustments in referral behavior. In the baseline specification (panel 5a), the relocation year exhibits a sharp and statistically significant jump in $\phi_{t'}$, followed by persistently higher coefficients. This pattern suggests that caseworkers quickly and durably adjust their broader referral behavior to the practice environment of the destination office. The persistence of post-move effects also indicates that these responses are unlikely to be driven by slow-moving mechanisms such as skill accumulation or learning (Molitor, 2018).

To address the reflection problem and isolate genuine peer influences from simultaneity and exogenous shocks from peers, I re-estimate the model, in column 1 and 2 of table 2, using time-invariant practice-environment constructed from pre-move differences between origin and destination offices. Because these office-level practices are defined prior to the mover's arrival, this specification avoids contemporaneous feedback and relies solely on variation plausibly exogenous to the caseworker's own behavior. The results show consistently positive and statistically significant spillover effects. Using pre-move practices, the estimated coefficients range from 0.438 to 0.257 across specifications, indicating that exposure to existing norms in the destination office has a meaningful influence on caseworker behavior and the findings persist even after controlling for local labor market shocks through LLM-by-year fixed effects.

Column 3 and 4 of table 2 report estimates that additionally control for differences in peer experience between origin and destination offices. The

Table 2. *Robustness of the spillover coefficient*

	Pre-move (Δ_n)		Δ in Peer Tenure		
	(1) Baseline	(2) Alternative	(3) Baseline	(4) Alternative	(5) Movers vs. non-movers
Spillover coeff. (ϕ)	0.438***	0.257**	0.443***	0.265**	0.407***
	(0.095)	(0.104)	(0.095)	(0.104)	(0.050)
Changes in peer tenure			-0.007*	-0.006*	
			(0.004)	(0.004)	
Mover dummy					0.003
					(0.004)
Covariates	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Caseworker FE	✓	✓	✓	✓	×
Local Labor Market × Year FE	×	✓	×	✓	×
Origin office FE	×	×	×	×	✓
Destination office FE	×	×	×	×	✓
Observations	112,713	111,324	115,870	114,470	626,248
R-squared	0.055	0.068	0.056	0.068	0.027

Notes: The table presents the difference-in-differences estimates based on the sample data from 2004-2018. Columns (1)–(2) use pre-move, time-invariant office practices to avoid simultaneity. Columns (3)–(4) additionally control for differences in average peer tenure across offices. Column (5) compares movers and non-movers. All regressions include job seeker covariates and standard errors clustered at the caseworker level. Asterisks ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

rationale is that if more experienced peers exert differential influence, variation in average peer tenure could account for part of the observed spillover. However, the coefficient on peer-tenure differences is extremely small and close to zero, and the spillover estimate remains stable and statistically significant. This robustness indicates that the spillover effects persist even after accounting for observable differences in peer characteristics, reinforcing the interpretation that the results reflect genuine behavioral adaptation rather than correlated traits. Next, I examine whether migrant and non-migrant caseworkers are systematically different in their broader referral decisions. Column 5 of table 2 compares broader referral behavior of movers and stayers within the same origin and destination office pairs within a given year. The coefficient of mover dummy is small and statistically insignificant, and the spillover coefficient remains strongly positive. This indicates that movers and stayers behave similarly when exposed to comparable office environments, providing evidence against systematic sorting by caseworker type. I further assess whether the two groups differ along other observable work-related dimensions, including referral frequency, meeting intensity, and caseload size. Appendix table A1 shows no systematic differences within office-year cells, indicating that movers are not selectively sorted on key aspects of their work practices. Together, these

findings strengthen the validity of the causal interpretation by ruling out sorting on observable caseworker or client characteristics. [17]

Heterogeneity in spillovers

Table 3. *Heterogeneity of spillover effects*

	Transition type		Caseworker experience	
	(1) Low→High	(2) High→Low	(3) High exp.	(4) Low exp.
Spillover coeff. (ϕ)	0.309***	0.294***	0.374***	0.286***
	(0.082)	(0.084)	(0.070)	(0.102)
Covariates	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Caseworker FE	✓	✓	✓	✓
Observations	59,161	56,624	72,578	43,207
R-squared	0.058	0.054	0.050	0.067

Notes: This table reports difference-in-differences (DiD) estimates using a time-variant measure of changes in the referral environment (Δ_{nt}) based on the sample data from 2004-2018. Columns (1)–(2) report results by transition type (low→high and high→low offices), and columns (3)–(4) by caseworker experience level (above- and below-median tenure). All regressions include year and caseworker fixed effects and control for job seeker characteristics. Standard errors are clustered at the caseworker level (in parentheses). Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

I examine heterogeneity by direction of transition (low→high vs. high→low) and caseworker experience. The estimates reported in column (1)-(2) of table 3 show that the magnitude and significance of the spillover coefficient are highly comparable across these two groups. This indicates that behavioral adjustments are not asymmetric: caseworkers tend to align their referral behavior with prevailing local norms regardless of whether they move into more or less intensive referral environments. The absence of directional asymmetry suggests that adaptation occurs broadly and consistently, reflecting an internalization of institutional norms rather than one-sided behavioral shifts. Finally, the results in column (3)-(4) also reveal no significant difference between caseworkers having different levels of experience. Both high- and low-experience caseworkers exhibit statistically significant and similarly sized spillover effects, implying that tenure does not affect the extent of behavioral adaptation, a result that is consistent with the estimates reported in column 3-4 of table 2. Taken together, these findings demonstrate that the behavioral spillover effects are stable, symmetric, and insensitive to specification choices, underscoring the

[17] For appendix table A1, I run the following simple regression using caseworker-year level sample data:

$$movers_dummy_n = \alpha + \Phi'_{nt}\beta + \zeta_{origin\,office} + + \zeta_{destination\,office} + \zeta_t + \varepsilon_{nt} \tag{6}$$

where Φ'_{nt} denotes the matrix of observable caseworker traits. I also control for origin office, destination office and year fixed effects to identify whether movers and stayers at the same office are different along these characteristics.

robustness and generality of the estimated influence of the referral environment on individual caseworker behavior.

Decomposing the spillover effects

To clarify the sources of the behavioral spillovers identified in the baseline specification, I decompose the estimated effect into peer-driven and office-specific components. The baseline DiD estimates capture changes in referral behavior following a move, but the aggregated practice-environment measure may combine peer-group influences with institutional or spatial office characteristics. The goal of this decomposition is therefore to isolate the portion of the spillover driven specifically by peers. To accomplish this, I include office fixed effects, which absorb all time-invariant office characteristics (e.g., managerial style, organizational routines, structural constraints) while leaving intact variation in peer-group composition. This design purges persistent office-level heterogeneity, ensuring that the remaining variation in the spillover coefficient reflects behavioral transmission among peers rather than confounding institutional conditions.

Table 4. *Decomposition of overall spillover effects using time-variant* Δ_{nt}

	(1)	(2)
Spillover coefficient (ϕ)	0.346***	0.284***
	(0.058)	(0.070)
Covariates	✓	✓
Year FE	✓	✓
Caseworker FE	✓	✓
Office FE	×	✓
Observations	115,785	115,781
R-squared	0.056	0.059

Notes: This table presents a decomposition of the spillover effects using a time-variant measure of the practice environment (Δ_{nt}) based on the sample data from 2004-2018. Column (1) corresponds to the baseline specification including year and caseworker fixed effects. Column (2) adds office fixed effects to isolate peer-group-specific spillovers by removing time-invariant office-level heterogeneity. All regressions are limited to migrant caseworkers and include covariates capturing job seeker characteristics. Standard errors, clustered at the caseworker level, are reported in parentheses. Asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4 presents the results. Column (1) shows the baseline spillover estimate of 0.346 using the time-variant measure of changes in the practice environment. When office fixed effects are added in column (2), the coefficient declines to 0.284. This difference quantifies the portion of the overall spillover attributable to stable office-specific factors. The estimates indicate that roughly four-fifths of the behavioral response arises from peer influence itself, while approximately one-fifth of the total spillover reflects time-invariant institutional features. In

other words, the referral behavior of coworkers, other than office-specific factors, accounts for the dominant share of the observed effect.

This pattern stands in contrast to settings such as healthcare, where institutional resources and infrastructure play a central role in shaping treatment behavior (e.g., Molitor 2018; Avdic et al. 2023). In the context of employment services, the modest office-specific component suggests that the physical or administrative environment exerts a limited role in determining referral practices. Instead, the strong peer-group component implies that caseworkers adapt primarily through social learning or assimilation to prevailing behavioral norms within their immediate team. This decomposition thus reinforces the interpretation that the spillover effects among caseworkers might be more interpersonal rather than institutional in nature, highlighting the centrality of peer norms in shaping broader referral behavior of caseworkers at the public employment service offices.

Effects on Job Seeker Outcomes

This section examines whether behavioral spillovers among migrant caseworkers translate into measurable labor market consequences for the job seekers they serve. The analysis focuses on two outcome dimensions: (i) unemployment exit probabilities (within 90 and 180 days), and (ii) post-unemployment earnings up to two years after reemployment. These outcomes could serve as meaningful indicators of the effectiveness of job placement strategies.

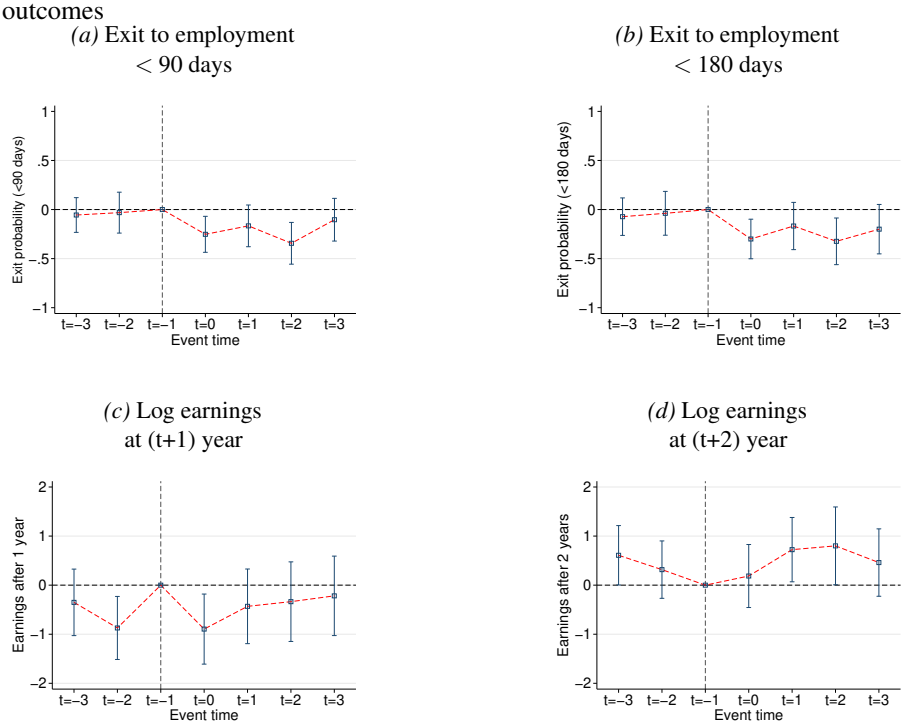
Table 5. *Effect of Broader Referral Environment on Job Seekers' Actual and Predicted Outcomes*

	Actual Outcomes				**Predicted Outcomes**	
	(1) Unemp. dur. < 90 days	(2) Unemp. dur. < 180 days	(3) Earnings at $t+1$	(4) Earnings at $t+2$	(5) Unemp. dur. < 90 days	(6) Unemp. dur. < 180 days
Change in practice-env. (Δ_{nt})	-0.138**	-0.150**	-0.173	-0.046	-0.003	0.000
	(0.055)	(0.066)	(0.172)	(0.147)	(0.003)	(0.008)
Covariates	✓	✓	✓	✓	×	×
Year FE	✓	✓	✓	✓	✓	✓
Caseworker FE	✓	✓	✓	✓	✓	✓
Local Labor Market × Year FE	✓	✓	✓	✓	✓	✓
Observations	114,385	114,385	76,129	76,105	114,385	114,385
R-squared	0.074	0.107	0.143	0.149	0.017	0.026

Notes: This table presents difference-in-difference estimates of changes in unemployment outcomes related to broader referrals based on the sample data from 2004-2018. "Change in practice-environment (Δ_{nt})" refers to the effect of the changes in the migrant caseworker's broader referral environment on their job seekers' unemployment exit probabilities (columns 1–2), future log earnings (columns 3–4), and a placebo duration outcome based on predicted probabilities of exit (columns 5–6). Each column is estimated on a movers-only sample using spell-level data. All specifications include year, caseworker, and local labor market-by-year fixed effects. All columns excluding the predicted outcomes (columns 5–6) control for job seekers' characteristics. Standard errors are clustered at the caseworker level and shown in parentheses. Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 5 presents the estimated effects of exposure to broader referral environments on unemployment exit probabilities. The results show that a 10 pp increase in a caseworker's exposure to broader referrals reduces the likelihood of her job seeker's unemployment exit within 90 days by approximately 1.4 pp, and within 180 days by roughly 1.5 pp respectively. Both effects are statistically significant at the 5 percent level. The magnitude and direction of these estimates suggest that when caseworkers adopt broader referral patterns, potentially due to peer conformity in offices with higher referral intensity, the matching process becomes less effective. One plausible explanation is that broader referrals may increase the number of job applications but simultaneously reduce job–person fit, leading to slower exits from unemployment. The results for post-unemployment earnings reveal a similar but less precise pattern. The estimated coefficients for earnings at $t+1$ and $t+2$ are insignificant.

Figure 6. Changes in practice-environment for caseworkers (Δ_n) and job seekers' outcomes



Notes: This figure is based on the sample data for years 2004–2018. This figure shows the the event-study graph of job seeker's duration and earnings outcomes in response to changes in caseworker's practice-environment. All panels show the responses with time-variant changes in practice-environment using caseworker, year, and interacted local labor market and year fixed effects. Standard errors are clustered at the caseworker level.

The dynamic pattern of these effects, illustrated in figure 6, further supports this interpretation. Following the move, exit probabilities decline persistently,

indicating that the behavioral adjustment of caseworkers to new peer environments has lasting implications for client outcomes. The parallel pre-trends reinforce the validity of the identification strategy. Overall, the findings imply that adaptation of caseworkers to a high-intensity broader referral environment may inadvertently undermine placement effectiveness in the short run. Broader referral practices appear to increase search inefficiency by encouraging job applications that might be poorly aligned with individual qualifications or constraints in this case.

One key concern in interpreting these results is potential selection bias in the job seekers assigned to migrant caseworkers before and after their move. If caseworkers systematically transition between offices serving different client populations (e.g., those with higher or lower employability), the observed outcome changes could reflect differences in job seeker characteristics rather than true behavioral effects. To address this concern, a placebo analysis was conducted using predicted exit probabilities, as estimated from a model based solely on observable pre-treatment characteristics (e.g., age, education, work history, prior unemployment, and demographics) of the job seekers. These predicted outcomes capture job seekers' baseline employability independent of caseworker behavior. As shown in table 5 (column 5 and 6), the estimated coefficients on the key treatment variable are small and statistically insignificant across all specifications, indicating the pool of job seekers are similar across origin- and destination-offices.

Another potential concern is the channel of effects. One might worry that changes in the practice environment may simultaneously induce changes in other observable aspects of caseworkers' behavior that are related to their tasks. To examine this, I estimate a simple model regressing the change in the practice environment (Δ_{nt}) on several observable caseworker-level characteristics[18]: the number of referrals per caseworker per year, the number of job seekers assigned to each caseworker per year, and the average number of meetings per job seeker per year. The results, shown in appendix table A2, indicate that other caseworker functions at the PES office are not associated with the changes in practice-environment. This finding suggests that the primary channel through which job seekers are influenced might be the broader referral environment itself, rather than changes in the frequency of referrals, the intensity of interactions with job seekers, or the overall caseload of the caseworkers.

[18] For this, I run the following regression using caseworker-year level data on movers sample:

$$\Delta_{nt} = \alpha + \Phi'_{nt}\beta + \zeta_n + \zeta_t + \varepsilon_{nt} \tag{7}$$

where Φ'_{nt} denotes the matrix of observable caseworker traits. I also control for caseworker and year fixed effects.

Heterogeneity in job seeker's outcomes

To examine whether the effects of broader referral environments vary across job seeker's characteristics, I estimate separate regressions by gender, age, and employability or quality. Table 6 reports the corresponding difference-in-differences estimates for unemployment exit probabilities and post-unemployment earnings (up to two years), with panels A–B covering gender, panels C–D covering age groups, and panels E-F covering the quality of job seekers based on predicted duration.

For male job seekers (Panel A), their caseworker's exposure to broader referral environment slightly reduces the exit probabilities within 90 and 180 days, but the estimates are small and statistically insignificant. In contrast, female job seekers (Panel B) exhibit larger and significant declines: a 10 pp increase in broader referral exposure to their caseworkers lowers the likelihood of exiting unemployment within 90 days by roughly 2.0 pp and within 180 days by 1.8 pp. These findings suggest that broad referral practices may disproportionately hinder women, possibly reflecting greater mismatches across broader occupations, gender differences in occupational sorting, or job preferences that increase the likelihood of mismatch. The post-unemployment earnings estimates are statistically insignificant, suggesting no earnings penalty.

Age heterogeneity, on the other hand, shows that caseworker's broader referral exposure reduces the exit probabilities of their younger job seekers (under 35) by about 2.2 pp for both exit horizons, with effects significant at the 1–5% levels (Panel C). For individuals aged 35 and above (Panel D), the estimates are near zero and statistically insignificant. Post-unemployment earnings effects remain insignificant across age groups. This suggests that younger job seekers are more affected by the inefficiencies of broad referrals, whereas older individuals, possibly with more stable work histories, are largely insulated from such effects.

Finally, Panels E–F show that the estimated effects are different for job seekers having different baseline exit risks. In particular, the adverse outcomes are driven almost entirely by individuals with lower predicted unemployment duration, job seekers who would otherwise be expected to return to employment relatively quickly. For this group, a 10 pp increase in caseworker exposure to broader referral environments reduces the probability of exiting unemployment within 90 days by approximately 2.1 pp, and within 180 days by about 2.3 pp. By contrast, individuals with higher predicted unemployment duration exhibit small and statistically insignificant responses. One explanation might be the fact that low-risk individuals are closer to the margin of efficient matching and therefore have more to lose from imprecise or generic referrals. For high-risk job seekers, limited opportunities and weaker labor market attachment may constrain the scope for referral-induced inefficiencies to matter. As a result, broader referral practices primarily disrupt efficient matches rather than exacerbating already poor prospects. Post-unemployment earnings remain

unaffected across both groups, suggesting that the effects operate through delayed reemployment rather than reduced job quality.

Table 6. *Heterogeneity in job seeker outcomes by gender, age, and predicted duration*

	Pr(Exit $<$ 90 days)	Pr(Exit $<$ 180 days)	Log(Earnings t+1)	Log(Earnings t+2)
Panel A: Males				
Change in practice-env. (Δ_{nt})	-0.084	-0.117	-0.101	0.024
	(0.068)	(0.081)	(0.239)	(0.209)
Observations	60,304	60,304	40,107	40,283
R-squared	0.090	0.121	0.170	0.168
Panel B: Females				
Change in practice-env. (Δ_{nt})	-0.196***	-0.178**	-0.372	-0.159
	(0.067)	(0.078)	(0.233)	(0.213)
Observations	53,787	53,787	35,719	35,509
R-squared	0.095	0.130	0.158	0.169
Panel C: Age <35				
Change in practice-env. (Δ_{nt})	-0.224***	-0.221**	0.004	-0.171
	(0.083)	(0.098)	(0.233)	(0.198)
Observations	65,313	65,313	43,781	43,492
R-squared	0.081	0.107	0.158	0.158
Panel D: Age ≥ 35				
Change in practice-env. (Δ_{nt})	-0.009	-0.013	-0.222	0.152
	(0.058)	(0.067)	(0.260)	(0.236)
Observations	48,753	48,753	32,008	32,270
R-squared	0.079	0.101	0.169	0.177
Panel E: Higher pred. duration				
Change in practice-env. (Δ_{nt})	-0.058	-0.069	-0.512*	0.070
	(0.048)	(0.065)	(0.265)	(0.236)
Observations	57,869	57,869	32,781	32,696
R-squared	0.067	0.082	0.145	0.151
Panel F: Lower pred. duration				
Change in practice-env. (Δ_{nt})	-0.213**	-0.233**	0.177	-0.125
	(0.088)	(0.107)	(0.221)	(0.197)
Observations	56,240	56,240	43,024	43,096
R-squared	0.071	0.087	0.192	0.191
Covariates	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Caseworker FE	✓	✓	✓	✓
Local Labor Market $\times$ Year FE	✓	✓	✓	✓

Notes: This table reports difference-in-differences estimates of the effects of broader referral environments on job seeker outcomes based on the sample data from 2004-2018, estimated separately by gender (Panels A–B), age group (Panels C–D), and predicted unemployment duration (Panels E–F). The treatment variable measures changes in the referral environment faced by migrant caseworkers following inter-office moves. All specifications include covariates, year, caseworker, and local labor market-by-year fixed effects. Robust standard errors clustered at the caseworker level are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Overall, the heterogeneity results suggest that when caseworkers adopt broader and more standardized referral strategies, potentially as a result of peer exposure, these approaches may crowd out caseworker discretion, leading to greater job–person mismatch and slower re-entry among some subgroup of workers who would otherwise be expected to match quickly. These findings underscore the importance of tailoring referral strategies to the characteristics

of job seekers to enhance matching efficiency and mitigate short-term inefficiencies introduced by peer-driven behavioral conformity among caseworkers.

Conclusion

This paper provides the first empirical evidence on how organizational environments shape caseworkers' vacancy referral behavior and, in turn, influence job seeker outcomes within public employment services. Using quasi-experimental variation from caseworker relocations across Swedish PES offices, I show that caseworkers systematically adjust their broader referral behavior to match the prevailing norms of their new peers. Difference-in-differences and event-study estimates reveal sharp behavioral shifts at relocation with no anticipatory trends, and these adjustments persist over time. Peer effects emerge as the primary mechanism, even after accounting for office-specific factors and local labor market conditions. The spillover operates symmetrically across transition types and does not vary by tenure, suggesting that conformity to local referral norms is a general feature of organizational adaptation. The magnitude of spillover coefficient also indicates the degree of consistency, in broader referral activity, with the prevailing norms of their destination office which is important from the organizational perspective.

These behavioral shifts have meaningful consequences for job seekers. Greater exposure to broader referral environments significantly lowers short-term reemployment probabilities, with a 10 pp increase in caseworker's exposure to broader referral practice at the new offices leads to the lower probability of exits (within 90 and 180 days) of their clients by about 1.4–1.5 pp. Although post-unemployment earnings remain unaffected, the results indicate that these frictions weaken early job matching rather than long-term job quality. Placebo tests using predicted outcomes confirm that the findings are not driven by selection or other caseworker functions, strengthening the causal interpretation. Heterogeneity analyses further indicate that these adverse effects are concentrated among women, younger job seekers, and those with lower predicted unemployment duration. This pattern suggests that broader referral practices are particularly disruptive to efficient matching for groups with relatively strong underlying reemployment prospects (e.g., young and workers with lower reemployment probability) and, for women, the effect may also reflect greater mismatch in applications to broader occupations, arising from gender differences in occupational sorting, job preferences, or the types of jobs considered acceptable or suitable.

Overall, the findings indicate that peer-induced conformity toward broader referral practices can introduce short-term mismatches. Although broader referrals are intended to expand opportunities, they may weaken the alignment between job seekers and suitable vacancies. This pattern aligns with Van der Klaauw and Vethaak (2022), who document the drawbacks of rigid search broadening, but contrasts with the results of Belot et al. (2019; 2022), who show

that flexible, information-based interventions can benefit highly disadvantaged workers. It is important to emphasize, however, that the comparison here is between more versus less broad referral environments, rather than evaluating broad referrals in isolation. The absence of long-term earnings penalties may suggest that the associated inefficiencies are transitory. More broadly, the results underscore the important role of peer norms and local institutional cultures in shaping frontline implementation of labor market policy. Designing referral systems that foster organizational learning while preserving client-centered discretion may help reduce mismatches and enhance the effectiveness of public employment services.

Declaration of Use of Generative AI

During the preparation of this manuscript, I used OpenAI's GPT-5.2 model to identify potential grammatical and typographical errors. All suggested corrections were carefully reviewed, and the manuscript was subsequently revised and finalized by me. I take full responsibility for the content of the article.

References

Altmann, S., Falk, A., Jäger, S., and Zimmermann, F. (2018). Learning about job search: A field experiment with job seekers in germany. *Journal of Public Economics*, 164:33–49.

Åslund, O. and Fredriksson, P. (2009). Peer effects in welfare dependence: quasi-experimental evidence. *Journal of human resources*, 44(3):798–825.

Avdic, D., Ivets, M., Lagerqvist, B., and Sriubaite, I. (2023). The evolution of physician practice styles: evidence from cardiologist migration. *Journal of Health Economics*, 89(102741):1–25.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, 77(4):1047–1094.

Belot, M., Kircher, P., and Muller, P. (2019). Providing advice to jobseekers at low cost: An experimental study on online advice. *The review of economic studies*, 86(4):1411–1447.

Belot, M., Kircher, P., and Muller, P. (2022). Do the long-term unemployed benefit from automated occupational advice during online job search? Technical report, IZA Discussion Papers.

Bollens, J. and Cockx, B. (2017). Effectiveness of a job vacancy referral scheme. *IZA Journal of Labor Policy*, 6:1–24.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):F452–F477.

Cederlöf, J., Söderström, M., and Vikström, J. (2021). What makes a good caseworker? Technical report, Working Paper.

Cheung, M., Egebark, J., Forslund, A., Laun, L., Rödin, M., and Vikström, J. (2024). Does job search assistance reduce unemployment? experimental evidence on displacement effects and mechanisms. Technical report, Working Paper.

Crépon, B., Dejemeppe, M., and Gurgand, M. (2005). Counseling the unemployed: does it lower unemployment duration and recurrence? *IZA discussion paper*.

Dhia, A. B., Crépon, B., Mbih, E., Paul-Delvaux, L., Picard, B., and Pons, V. (2022). Can a website bring unemployment down? experimental evidence from france. Technical report, National Bureau of Economic Research.

Hedin, G. and Tegsjö, B. (2006). Lokala arbetsmarknader: egenskaper, struktur och utveckling. Arbetsmarknadsstatistisk rapport, am0207, Statistiska centralbyrån (SCB), Stockholm.

Ichino, A. and Maggi, G. (2000). Work environment and individual background: Explaining regional shirking differentials in a large italian firm. *The Quarterly Journal of Economics*, 115(3):1057–1090.

Lagerström, J. (2011). How important are caseworkers-and why? new evidence from swedish employment offices. Technical report, Working Paper.

Lechner, M. and Smith, J. (2007). What is the value added by caseworkers? *Labour economics*, 14(2):135–151.

Lundin, D. (2004). Vad styr arbetsförmedlarna? Technical report, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Lundin, M. and Thelander, J. (2012). Ner och upp–decentralisering och centralisering inom svensk arbetsmarknadspolitik 1995–2010. Technical report, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–145.

Molitor, D. (2018). The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy*, 10(1):326–356.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, 116(2):681–704.

Schiprowski, A. (2020). The role of caseworkers in unemployment insurance: Evidence from unplanned absences. *Journal of Labor Economics*, 38(4):1189–1225.

Van den Berg, G. J., Kjærsgaard, L., and Rosholm, M. (2014). To meet or not to meet, that is the question: short-run effects of high-frequency meetings with case workers. Technical report, Working Paper.

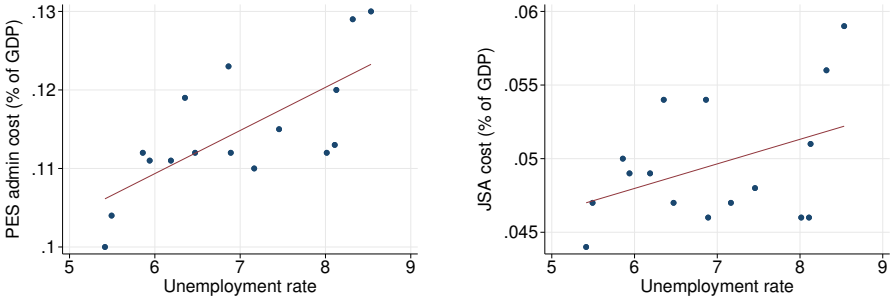
Van der Klaauw, B. and Vethaak, H. (2022). Empirical evaluation of broader job search requirements for unemployed workers. Technical report, JSTOR.

Weber, A. and Hofer, H. (2003). Active job-search programs a promising tool? a microeconometric evaluation for austria. Technical report, Reihe Ökonomie/Economics Series.

Appendix

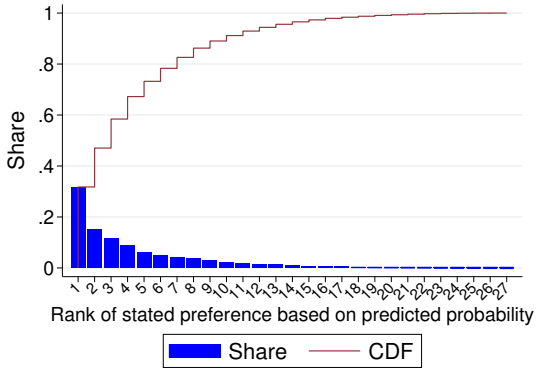
Descriptive statistics

Figure A1. Relationship between unemployment rate, public employment service spending, and JSA spending in European countries



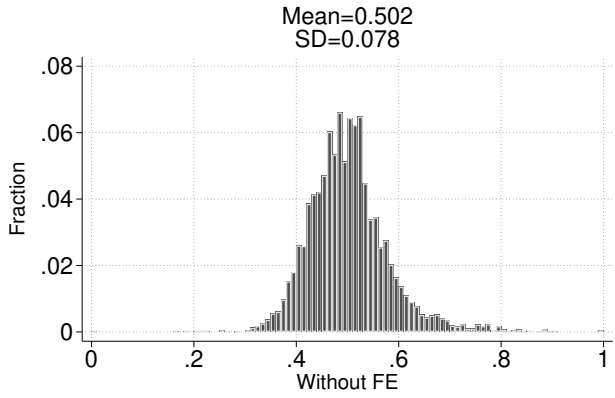
Notes: The figures present binscatter plots of OECD region-year observations showing the relationship between unemployment rates and (i) public employment service (PES) administrative expenditures and (ii) job-search assistance (JSA) expenditures, both expressed as a percentage of GDP using OECD country-year observations from 2005–2020. Each point represents the average value within equally sized bins of the unemployment rate distribution. The solid line indicates the fitted linear regression. The figures are descriptive and illustrate the positive association between unemployment and PES and JSA spending.

Figure A2. Are stated preference of job seekers relevant?



Notes: This figure shows the relationship between job seekers' stated and predicted preferences based on the sample data from 2004-2018. The x-axis ranks stated preferences by predicted probability of choosing that occupation by a job seeker conditional on job seekers characteristics, and the y-axis shows the share of job seekers. Blue bars depict the distribution of stated preferences, while the maroon line represents the empirical cumulative distribution function (CDF).

Figure A3. Practice Environment of Migrant Caseworkers



Notes: Based on the sample data from 2004–2018, figure A3 shows the distribution of practice environments across the offices to which migrant caseworkers are assigned. This represents a raw, cross-sectional distribution that captures differences in referral practices across offices. The wide dispersion highlights substantial within-caseworker variation, which is essential for empirical identification.

Table A1. *Are movers different than stayers?*

	(1) Mover's dummy
# of referrals per caseworker in a year	-0.000
	(0.000)
# of job seekers per caseworker per year	0.000
	(0.000)
Average # of meetings per job seeker	0.000
	(0.000)
Origin Office FE	✓
Destination Office FE	✓
Year FE	✓
Observations	36,876
R-squared	0.405

Notes: This table presents diagnostic regressions examining whether movers differ systematically in their observable characteristics from stayers using caseworker-year-level data from 2004-2018. The dependent variable is the mover dummy. Each coefficient represents the partial association between the characteristic and the likelihood of moving. All regressions include origin and destination office fixed effects as well as year fixed effects. Standard errors are clustered at the caseworker level and shown in parentheses. Asterisks ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Table A2. *Do changes in practice-environment affect migrant caseworkers' other traits?*

	(1) Δ_{nt}
# of referrals per caseworker in a year	0.000
	(0.000)
# of job seekers per caseworker per year	-0.000
	(0.000)
Average # of meetings per job seeker	0.000
	(0.000)
Caseworker FE	✓
Year FE	✓
Observations	6,605
R-squared	0.454

Notes: This table reports diagnostic regressions testing whether changes in the practice environment of movers affect other observable caseworker characteristics using caseworker-year-level data from 2004-2018. The dependent variable is the change in practice environment (Δ_{nt}). All specifications include caseworker and year fixed effects. Standard errors are clustered at the caseworker level and shown in parentheses. Asterisks ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Returns to Breaking the Norm: Evidence from Job Referrals[1]

Co-authored with Jonas Cederlöf and Johan Vikström

[1]The authors would like to thank Peter Fredriksson, Björn Öckert, Anna Sjögren, Yoko Okuyama, Georg Graetz, Olof Åslund, Caroline Hall, Adrian Adermon, Lena Hensvik as well participants at UCLS Annual Workshop, IFAU Seminar Conference, and Uppsala Labor Group.

Introduction

Occupational segregation remains a central source of gender inequality in the labor market. Despite the sharp rise in female labor force participation and substantial occupational integration over the twentieth century, progress toward gender-balanced occupations has slowed markedly since the early 2000s (Blau et al., 2013). A large body of research has shown that persistent and misperceived gender norms continue to shape occupational choices, helping sustain segregation across jobs (Cejka and Eagly, 1999; Cortes and Pan, 2018; Cortés et al., 2026; Bursztyn et al., 2024; Del Carpio and Guadalupe, 2022; Barbulescu and Bidwell, 2013; Delfino, 2024; Sorkin, 2017; Alesina et al., 2013) This matters because male-dominated occupations typically offer higher wages and better career prospects, making occupation and workplace segregation an important contributor to the gender wage gap (Blau and Kahn, 2017; Card et al., 2015; Bamieh and Ziegler, 2024). World Bank (2023) estimates suggest that occupational segregation accounted for 33% of the gender wage gap in the United States in 2010, exceeding the contributions of industry (18%) and experience (14%).[2] Beyond its distributional consequences, segregation may also lower efficiency by misallocating talent across occupations. A key question is then whether public policy can weaken the gendered beliefs and constraints that shape occupational choices and thereby reduce occupational segregation.

This paper studies how a specific type public policy can help break occupational gender segregation in the labor market. Namely, how caseworkers at the Public Employment Services (PES) can be used to influence and broaden unemployed job seekers' perceptions about what constitutes a "viable" job/occupation and how this in turn affects labor market outcomes. To this end, we exploit the fact that caseworkers refer their job seekers to different open vacancies, and infer caseworkers' non-stereotypical gender attitudes by their propensity to provide job seekers with vacancy referrals to non-stereotypical occupations. To account for the fact that caseworkers' referral behavior is endogenous to the set of assigned job seekers, we exploit the use of a date-of-birth rule to allocate job seekers to caseworkers. Since the exact day of birth of the job seeker is orthogonal to her characteristics, this mimics random assignment of caseworkers. To handle occasional exemptions from the date-of-birth rule, we use an IV-framework based on the caseworker each job seeker would have had, had she been allocated using the date-of-birth rule.

We start by documenting that also in Sweden, a highly gender egalitarian country, there exists substantial occupational segregation. Moreover, using data on unemployed job seekers' preferred occupations which are stated at the

[2]Interestingly, as documented by World Bank (2023) no occupation appears to be universally dominated by men or women; several occupations are male- or female-dominated in some countries but gender-balanced in others, highlighting the role of social and institutional factors and the scope for policy intervention.

onset of unemployment, we document strong gender patterns in job search. Both males and females are substantially more likely to seek employment in occupation which is dominated by their own gender. These gender difference remain when controlling for a rich set of background characteristics such as demographics, type of education, as well as other occupation characteristics such as median wages and hours worked. It even remains after adjusting for previous occupations and industries. While descriptive, these results indicate that both men and women are influenced by gender-norms in their occupational choices.

We then turn to how referrals made by caseworkers at the PES can be used as a policy to break gender-stereotypical occupational choices. To this end, we measure the non-stereotypicalness of each caseworker based on what type of referrals they give their assigned job seekers. As the type of referral caseworkers provide is endogenous to the set of job seekers they are assigned, we follow Cederlöf et al. (2025) and Humlum et al. (2025), and exploit as-if random assignment of job seekers to caseworkers based on a date-of-birth rule. To handle exemptions from this rule, we rely on an IV-strategy instrumenting caseworkers actual non-stereotypicalness with that of the caseworker to whom the job seekers would have been assigned according to the rule (the predicted caseworker). We verify that our instrument is predictive of actual caseworkers' referral behavior and validate its independence with respect to job seeker characteristics.

Our findings reveal that exposure to a more gender non-stereotypical caseworker improves subsequent labor market outcomes. A one standard deviation increase in caseworker non-stereotypicalness raises the probability of exiting unemployment within 180 days by 5.7 percentage points (corresponding to a 10% increase) and increases annual earnings by 4% in the year following unemployment. These earnings gains persist over time, although decline to about 2%. This suggests that encouraging job seekers to consider gender-atypical job opportunities appears to improve matching efficiency by expanding the set of jobs that job seekers consider and by relaxing inefficient gendered constraints in job search. This interpretation is reinforced when we interact the caseworker non-stereotypical behavior with the job seekers' occupational background. The effects are primarily driven by job seekers having previously worked in a gender-stereotypical occupation, consistent with non-stereotypical caseworkers being particularly valuable for job seekers whose prior careers suggest stronger gendered constraints in search.

We continue to study whether the effects of interacting with a non-stereotypical caseworker differentially affects male and female job seekers. Interestingly, we find substantial heterogeneity in this dimension with male job seekers having twice as large of an effect from meeting a non-stereotypical caseworker on the probability of leaving unemployment within 180 days. Moreover, they fully account for the positive earnings effect both in the short

and long run. We interpret this as expanding the set of viable job opportunities for male job seekers seem more important than for female job seeker.

As a caseworker's gender stereotypical behavior may very well correlate with other caseworker traits that potentially also affect job seekers' labor market outcomes, we probe the robustness of our results by measuring and controlling for other caseworker traits. Our findings are remarkably robust to the inclusion of e.g., caseworker tenure, meeting intensity, referral volume and propensity to assign job seekers to labor market programs. However, we find that part of the effect of caseworkers' non-stereotypicalness appears to reflect broader job search in general and referrals towards tighter occupations.

In the final part of the paper we turn to mechanisms and show that non-stereotypical caseworkers also affect the kinds of jobs workers obtain. For women, exposure to a more non-stereotypical caseworker increases transitions into male-dominated occupations and raises wages in the first post-unemployment job, suggesting that such caseworkers can help women access better-paid occupational paths. For men, we find no corresponding effect on occupational destination or wages, despite clear gains in earnings overall, implying that their benefits mainly operate through improved employment prospects rather than higher wages conditional on employment. Taken together, these findings suggest that non-stereotypical caseworkers influence not only whether job seekers find work, but also, at least for some groups, the type of work they enter.

Our paper relates and contribute to several strands of literature: First, and most closely related are Pan (2015) and Delfino (2024) who explore barriers to men's entry into female-dominated fields, finding that perceived gender shares have limited impact on application rates. Barbulescu and Bidwell (2013) demonstrate how women's job applications are shaped by preferences for work-life balance and perceptions of success in male-dominated roles. Second, we contribute to research on gender differences in job search behavior. Prior studies document persistent gender gaps in applications and occupational choices even among observationally similar workers (Barbulescu and Bidwell, 2013; Cortés et al., 2023; Fluchtmann et al., 2024), where we show that manipulating perceived gender composition has limited effects on men's applications to female-dominated fields (Pan, 2015; Delfino, 2024). Third, we add to the growing literature studying the influence of caseworkers on job seekers subsequent labor market outcomes through various channels (Schiprowski, 2020; Cederlöf et al., 2025; Humlum et al., 2025). Finally, we relate to the literature on how policy can be used to combat occupation and workplace segregation (see e.g., Card et al., 2025).

The remainder of the paper proceeds as follows. The next sections document occupational gender segregation in Sweden and show that unemployed job seekers exhibit strongly gender-segregated occupational preferences, before describing the institutional setting of the Swedish Public Employment Service, the role of caseworkers, and our measure of caseworker gender non-

stereotypicalness. The paper then presents the data and descriptive statistics, outlines the empirical strategy and evidence supporting the validity of the identification design, and reports the main effects of exposure to more gender non-stereotypical caseworkers on unemployment exits and subsequent earnings, including heterogeneity by prior occupational background and gender. The penultimate section examines underlying mechanisms through occupational transitions and wages in post-unemployment jobs. The final section concludes.

Background: Gender-segregated labor-markets

Occupational segregation in the US and Sweden

Despite some improvements during the postwar period, occupations remain gender segregated in the US (Blau et al., 2013; Hegewisch et al., 2010; England et al., 2020; Blau, 2025; Liepmann and Hegewisch, 2025). Female-dominated occupations are common in education, healthcare and administrative services, while men dominate in construction, manufacturing and technical occupations. One reason this is relevant is that it contributes to gender wage-gaps (Blau and Kahn, 2017; Hegewisch and Hartmann, 2014; Olivetti et al., 2024; Foster et al., 2020). The same phenomena are present in most countries.

Figure 1a shows that gender segregation also is an issue in Sweden.[3] Panel A shows the distribution of the share of female workers across occupations. A substantial share of the occupations remains concentrated at the extremes, highly dominated by male or female workers, but there are also quite some occupations with a more balanced gender composition. Panel B, which plots the average level of occupational segregation[4] over time, shows that this segregation is highly persistent, even if there is a very modest trend towards less segregation. Here, a segregation value of 0.4 implies that the difference between male and female employment shares across occupations, on average, is 40 percentage points, reflecting a high degree of gender segregation in the labor market.

[3]The statistics are based on 3-digit occupational data, constructed using the wage survey data described in Section for 2014–2017.

[4]Here, occupational segregation is measured as the mean absolute difference between male and female employment shares across occupations in a given year, computed as

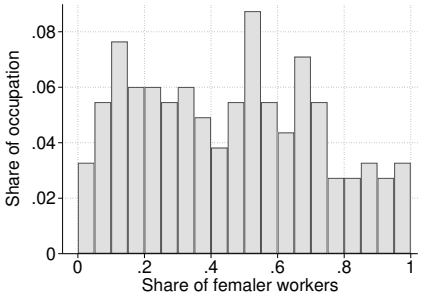
$$\frac{1}{N}\sum_{i=1}^{N}\left|\frac{m_i}{M}-\frac{f_i}{F}\right|,$$

where m_i and f_i denote male and female employment in occupation i, M and F denote total male and female employment, and N is the number of occupational categories. This measure can be interpreted as a normalized variant of the Duncan dissimilarity index, capturing the average per-occupation gender imbalance rather than the proportion of workers who would need to change occupations.

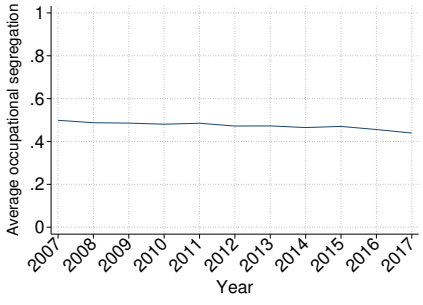
In this paper, we will examine male-dominated occupations with more than 60 percent males in the occupation, with an analogous definition of female-dominated occupations. Appendix table A1 shows that most male-dominated occupations are construction, mechanical repair, manufacturing, and the military, in which, there is well below 10% females. The most female-dominated occupations are office secretaries, preschool teachers, nurses, and midwives, with more than 85% females. When examining trends, appendix figure A1 shows that the mild shift towards less segregation mainly is due to a decline in the number of male-dominated occupations.

Figure 1. Gender composition and occupational segregation in Sweden

(a) Distribution of female shares across occupations



(b) Occupational segregation over time



Notes: Panel (a) shows the distribution of female employment shares across 3-digit occupations in Sweden, computed using annual wage survey data from Statistics Sweden for the period 2014–2017. Female shares are calculated by pooling observations across years and dividing the number of female workers by total employment within each occupation. Panel (b) plots average occupational segregation over time, measured as $\frac{1}{N}\sum_{i=1}^{N}\left|\frac{m_i}{M}-\frac{f_i}{F}\right|$, where m_i and f_i denote male and female employment in occupation i, and M and F denote total male and female employment in the labor force at time t. Occupational segregation is estimated using annual wage survey data from 2007 to 2017 at the 3-digit occupational level. The solid line represents the average segregation index across occupations.

Segregated job-search

We next examine how gender-based occupational segregation influences the job search behavior of male and female workers. To this end, we exploit the fact that all Swedish workers who become unemployed and register with the Public Employment Service (PES) are required to report an occupation on which they intend to primarily focus their job search[5]. We then investigate whether male (female) workers prefer male- (female-) dominated occupations to a larger extent and examine the contributing factors.

[5]Unemployed workers report up to four occupations (73% state more than one occupation), but here we only use the first, noting that we obtain similar results if we consider all four.

Panel A of table 1 presents the results from regressing an indicator for a preferred female-dominated occupation on a female worker dummy. It reveals striking gender patterns also in job-search behavior. Female workers are 52.7 percentage points more likely to focus their search on a female-dominated occupation compared to males (column 1). This drops to 51.1 percentage points and to 38.0 percentage points when adjusting for demographics and detailed measures of type of education, respectively (columns 2–3). It further drops to 13.7 percentage points when we adjust for previous occupation and industry fixed effects (columns 5–6). Panel B reveals a corresponding pattern for male workers and male-dominated occupations.

It shows that gender segregation is substantial for males and females with the same education, and a substantial gender job-search gap even exists if we compare females and males who worked in the same occupation before becoming unemployed. That said, previous occupation is still a strong predictor of the direction of search. In sum, we conclude that gender segregation very much shapes job search and without intervention from policy, job seekers appear to contribute to persistent gender segregation.

Table 1. *Do job seekers focus their job-search on gender-dominated occupations?*

	Panel A: Search for a job in *female-dominated occupation*					
	(1)	(2)	(3)	(4)	(5)	(6)
Female dummy	0.527***	0.511***	0.380***	0.192***	0.140***	0.137***
	(0.018)	(0.014)	(0.009)	(0.007)	(0.005)	(0.005)
	Panel B: Search for a job in *male-dominated occupation*					
	(1)	(2)	(3)	(4)	(5)	(6)
Male dummy	0.525***	0.492***	0.350***	0.182***	0.126***	0.122***
	(0.020)	(0.015)	(0.011)	(0.006)	(0.004)	(0.004)
Observations	1,499,686	1,499,686	1,499,686	1,499,686	1,499,686	1,499,686
Demographics	×	✓	✓	✓	✓	✓
Type of education	×	×	✓	✓	✓	✓
Preferred occupational characteristics	×	×	×	✓	✓	✓
Previous occupation fixed effects	×	×	×	×	✓	✓
Industry fixed effects	×	×	×	×	×	✓

Notes: The sample consists of unique unemployment spells in Sweden from 2007–2017. In Panel A, preference for a female-dominated occupation is regressed on the female dummy. In Panel B, preference for a male-dominated occupation is regressed on the male dummy. The dependent variable equals one if the job seeker's first stated occupational preference is a gender-dominated occupation and zero otherwise. Demographics include key individual characteristics. Types of education are measured using 3-digit education codes. Preferred occupational characteristics include median income and average time-share in the preferred occupation. Previous occupation fixed effects use 3-digit occupational codes and industry fixed effects use 2-digit industry codes. Standard errors are clustered at the regional (län) level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Institutional setting

Public employment service and caseworkers

Job search in Sweden is largely organized through the Public Employment Service (PES), which plays a central role in modern labor market policy by supporting unemployed individuals in their transition back to work. To remain eligible for unemployment benefits, individuals must register with the PES, actively search for work, regularly report their search activity, and accept suitable job offers or assigned labor market programs. The PES offers a variety of support, including job-search assistance, re-training and other labor market programs.

The PES is organized in around 300 local offices, providing support to all job seekers in the local area. Once a job seeker registers as unemployed, he or she is assigned a caseworker. Caseworkers are the key actors through which services are provided to the unemployed. They hold meetings with job seekers, assess employability, refer job seekers to vacant jobs, and determine participation in programs. Despite operating under common guidelines and regulations, caseworkers in Sweden typically enjoy substantial discretion in how they deliver services (Lundin, 2004; Lagerström, 2011). Two reasons are that the guidelines caseworkers should follow are vaguely formulated, and caseworkers are evaluated based on how quickly job seekers find jobs rather than the specific programs and strategies they employ.

Stereotypical and non-stereotypical caseworkers

This has several implications for this study. First, caseworkers decide how frequently to meet with their job seekers and have considerable discretion over the support they provide during the meetings. One central objective is to provide information about, and refer job seekers to, suitable jobs. Thus, caseworkers' ideas about what constitutes suitable jobs and the vacancy referral they communicate to their job seekers affect the type of support job seekers receive, and caseworkers differ greatly in their propensity to refer job seekers to a male- or a female-dominated occupation. We use this variation in the types of vacancy referrals caseworkers send to classify them as stereotypical or non-stereotypical. We use data on all vacancy referrals sent by the caseworkers, and define a referral as non-stereotypical if a male job seeker is referred to a female-dominated occupation (as above, more than 60% females in the occupation), and vice versa. We then use the share of non-stereotypical referrals for each caseworker as a measure of how non-stereotypical a caseworker is.

Second, caseworkers' perceptions of occupational segregation and suitable jobs may also affect other actions they take. For instance, these perceptions may influence how they interact with male and female job seekers during meetings, which in turn may affect job finding. We therefore consider our

non-stereotypical measure to reflect a broader caseworker trait, rather than something exclusively confined to vacancy referrals. Thus, we will measure the reduced form effect of a non-stereotypical caseworker, and not attempt to estimate the isolated effect of a non-stereotypical referral.

Third, caseworkers may differ along other dimensions besides stereotypicalness, such as how frequently they meet with their job seekers and how they assign job seekers to labor market programs. Since these non-gendered caseworker characteristics may be correlated with stereotypicalness, we perform various robustness analyses using detailed data on such alternative strategies.

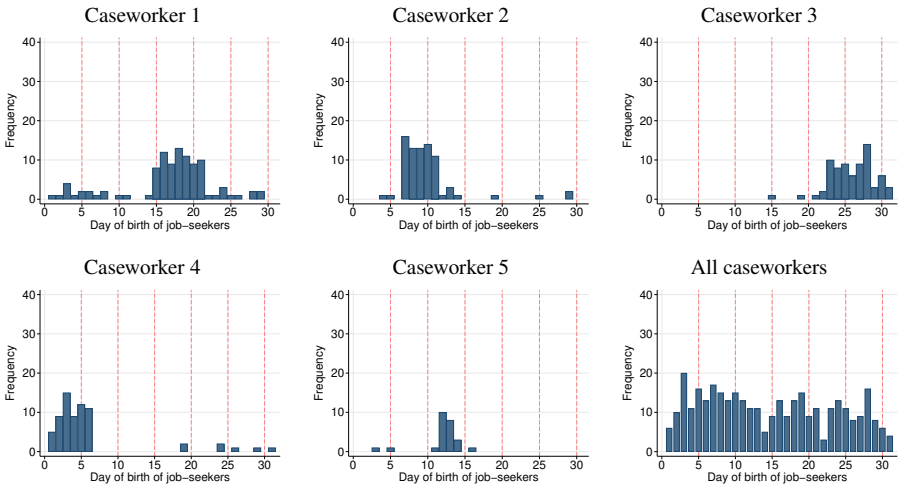
Fourth, promoting search in non-stereotypical occupations implicitly also promotes search in other types of occupations, both by broadening the search and potentially directing it toward tighter occupations with more available jobs. To examine this, we construct information on caseworkers behavior also in these dimension.

Date-of-birth assignment

A key challenge when studying the impacts of a non-stereotypical caseworker is non-random assignment of job seekers to caseworkers. In many institutional settings, job seekers are systematically sorted based on observed or unobserved characteristics such as employability, prior labor market attachment, or special needs, which confounds the identification of caseworker effects. Following Cederlöf et al. (2025) and Humlum et al. (2025), we address this challenge by exploiting an institutional feature used by several local employment offices in Sweden: assignment of job seekers to caseworkers based on their date of birth. Under this rule, caseworkers are responsible for job seekers born within specific intervals of the month. For example, one caseworker may handle job seekers born on the 1^{st} to 5^{th} of each month, another those born on the 6^{th} to 10^{th}, and so on. Since the exact day of birth is unrelated to job seekers' labor market characteristics, this creates as-if random variation in caseworker assignment within an employment office, effectively randomizing whether job seekers get a stereotypical or non-stereotypical caseworker.

This is illustrated in figure 2, which plots the distribution of job seekers born on each day of the month for an office with such a date-of-birth rule. The pronounced clustering of job seekers' birth dates within specific intervals for the caseworkers exemplifies the date-of-birth rule. However, the figure also reveals exemptions from the rule as the caseworkers also have job seekers born outside of their birth interval. One reason is that job seekers with special needs occasionally are reassigned to another caseworker. To account for these exemptions which may introduce non-random sorting, we adopt an instrumental variables approach, as in Cederlöf et al. (2025), based on a predicted caseworker. For each employment office, year, and day of the month, we define the predicted caseworker as the caseworker with the largest number of job seekers born on that specific day. Intuitively, this corresponds to the caseworker a job seeker

Figure 2. Date-of-Birth Randomization in Caseworker Assignment



Notes: The figure shows the allocation of job seekers to caseworkers by day of birth (1–31) in a local employment office applying a date-of-birth assignment rule. Each panel plots the frequency of job seekers by day of birth for a given caseworker. Vertical dashed lines indicate the cutoffs used to divide responsibility across caseworkers. The final panel aggregates all caseworkers within the office. The strong clustering of dates within specific intervals reflects systematic assignment based on date of birth, while observations outside these intervals indicate occasional exemptions from the rule.

would have been assigned to if the office had followed the date-of-birth rule strictly, without exemptions. We then use this predicted caseworker as an instrument for the actual caseworker assignment. This strategy isolates the random component of caseworker assignment induced by the date-of-birth rule.

These offices constitute a substantial share[6] of the Public Employment Service and are, on average, larger than other offices (more caseworker and job seekers), but in terms of job seeker characteristics they are similar to the offices without a date-of-birth rule, supporting the external validity of our results (appendix table A2).

Data & Descriptives

Data

Our main data source is the PES unemployment register encompassing the universe of unemployed job seekers registered between 2007–2015. These data contain for each job seeker the start and end date of their unemployment spell(s)

[6]An office's status as a date-of-birth office is not necessarily constant over the sample period, since some offices apply the date-of-birth assignment rule only during particular sub-periods. Although approximately 77 percent of unique offices in our sample use the rule at some point during the observation period, only around 39 percent of unemployment spells in the final sample belong to the office-year periods in which the rule is in force.

as well as daily information about participation in labor market programs, subsidized employment, meetings with caseworkers, etcetera. This register also contain the exact date of birth of each job seeker as well as occupational preferences of the job seeker stated at the onset of unemployment which are used to characterize job-search behavior and gendered occupational preferences. Importantly, through these registers we are able to link job seekers to every caseworker they have interacted with and whom made any type of decision regarding their job search process (e.g. a vacancy referral).

We combine these data with multiple population-wide registers from Statistics Sweden (SCB): (*i*) the Longitudinal Integration Database for Health Insurance and Labor Market Studies (LISA), where we obtain demographic information such as gender, age, country of birth, and level of education; (*ii*), matched employer-employee data (RAMS) containing information on annual earnings and monthly employment indicators; (*iii*) the Wage Survey comprises the employer-provided full-time-equivalent monthly wage, measured in September-November for all workers in the public sector and around half of all private-sector workers.

Finally, in a separate register linked to our main data source, we observe the universe of vacancy referrals made to job seekers, together with information on the caseworker issuing each referral and the date of issuance. Through a vacancy ID-number, we link all referrals to the PES online job-board (*Platsbanken*) which contain detailed information about each job-opening, such as occupation, industry and geographical location.

Descriptive statistics

Table 2 presents summary statistics at the office, job seeker, and caseworker levels in Panels (a), (b), and (c), respectively, for the date-of-birth (DOB) offices used in our analyses. The average number of caseworkers per office is 19. Job seekers are on average about 32 years old, slightly more than half are male, and educational attainment is broadly similar across office types, with roughly one third having at most pre-high school education and about one fifth holding some form of tertiary education. Around 17 percent are married, about 30 percent have children and more than half were employed one year prior to the unemployment spell. Turning to caseworker characteristics in Panel (c) we see that caseworkers make on average around 63 vacancy referrals per year and handle about 58 job seekers per year. The majority of caseworkers have substantial tenure at the PES: around 1 percent have less than two years of experience, while more than 8 percent have at least ten years of tenure.

Defining gender stereotypical referral behavior

As described earlier, we use these data on caseworkers' vacancy referrals to infer whether caseworkers' perceptions of what constitutes a suitable job can be classified as gender stereotypical or non-stereotypical, where we define an

Table 2. *Descriptive statistics for date-of-birth offices*

	Date-of-birth offices	
	Mean	SD
Panel (a): Office characteristics		
Number of caseworkers	18.981	11.995
Number of unique job seekers	860.752	573.177
Number of job seeker spells	1072.677	712.033
# Obs. (office-year panel)	517	
Panel (b): Job seeker characteristics		
Male	0.544	0.498
Age	31.883	12.840
Pre-high school $<$ 9 years	0.014	0.116
Pre-high school 9 years	0.288	0.453
High school 2 years	0.173	0.378
High school 3 years	0.311	0.463
University college 3 years	0.098	0.298
University college $\geq$ 3 years / PhD	0.110	0.313
Married	0.169	0.374
Children	0.308	0.462
Employed 1 year before spell	0.560	0.496
Any program in last 4 years	0.034	0.181
Number of unemployment spells (5 yrs)	1.390	1.890
# Obs. (spell level panel)	554,574	
Panel (c): Caseworker characteristics		
Referral behavior:		
Total referrals*	63.283	90.293
Share of non-stereotypical referrals*	0.212	0.179
Meetings and workload:		
Number of meetings (annual)	209.010	224.258
Number of job seekers per year	57.955	61.779
Share of job seekers assigned to ALMP	0.250	0.201
Tenure:		
0–2 years	0.010	0.099
2–5 years	0.382	0.486
5–10 years	0.526	0.499
10+ years	0.083	0.275
# Obs. (caseworker-year panel)	9,569	

Notes: The table reports means and standard deviations for office-, job seeker-, and caseworker-level characteristics in date-of-birth offices. Job seeker characteristics are measured at the start of the unemployment spell. '*' denotes that the caseworker-year panel includes only caseworkers who give at least one referral (here, # Obs. = 9,567).

occupation as gender-stereotypical if more than 60% of its workforce consists of one gender. Specifically, for each job seeker i, registered at office k and assigned to caseworker j, we calculate the share of gender non-stereotypical referrals given by caseworker j to all her clients N_j, in year t. To separate between caseworker's own behavior as rather than mechanically reflecting the

individual job seekers' preferences or suitability for a stereotypical vis-a-vis non-stereotypical occupation, we exclude job seeker i from the calculation and define a leave-out-mean of the form:

$$NS_{j(-i)kt} = \frac{1}{N_{(j-1)kt}} \sum_{m \neq i}^{N_{j-1}} \mathbf{1}[\text{non-stereotypical referral}]_{mkt}, \tag{8}$$

where $\mathbf{1}[\cdot]$ is the indicator function equal to one if the referral provided to job seeker $m \neq i$ is non-stereotypical. To minimize the influence of measurement error in the leave-out-mean, we restrict our sample to job seekers assigned to caseworkers having made at least 50 referrals. Figure A2 in the appendix shows the distribution of the share of average non-stereotypical referrals for our sample where higher values indicating less gender-stereotypical referral behavior. It reveals substantial heterogeneity in referral behavior across caseworkers; at the 25th percentile of the distribution, caseworkers the share of non-stereotypical referrals is around 19.5%, while the the corresponding number for caseworkers at the 75th percentile is just above 35%.

Naturally, what type of vacancies caseworkers recommend their clients to apply for will be endogenous to the type of job seekers they are assigned. For example, if a caseworker specializes in assisting job seekers within a particular industry and therefore is assigned all job seekers seeking employment in e.g., construction, NS will primarily reflect the which type of clients the caseworker handles—their preferences and suitability for a given job—and less so, the caseworker's own intrinsic/underlying gender stereotypical perceptions/behavior. In the next section, we describe in detail how we tackle this endogeneity by using the referral behavior of the *predicted caseworker* as an instrument for that of the actual caseworker.

Identification strategy

Our identification strategy relies on quasi-random variation in job seekers' exposure to caseworkers with different levels of non-stereotypicalness. Using the institutional assignment rule, we construct an instrument for the non-stereotypicalness of the assigned (actual) caseworker. This section first outlines the empirical strategy and the first-stage relationship generated by the assignment mechanism. It then discusses the assumptions underlying the IV design—independence, monotonicity, and the exclusion restriction—and argues why they are plausible in our context.

Empirical strategy

Our baseline empirical model is given by:

$$y_{ijkt} = \beta NS_{j(i)kt} + (\phi_k \times \gamma_t) + \delta X' + \varepsilon_{ijkt}, \tag{9}$$

where y_{ijkt} denotes the outcome (e.g. labor earnings) and denote $(\phi_k \times \gamma_t)$ office-by-year fixed effects which ensure that identification relies exclusively

on within-office, within-year variation in caseworker exposure. X' is a vector of pre-determined job seeker characteristics included primarily to increase precision. If job seekers were randomly allocated to caseworkers, β would reflect the causal effect of moving up one standard deviation in the gender non-stereotypical caseworker distribution.

Although we restrict our sample to local employment offices with a high degree of date-of-birth assignment, occasional deviations from the rule, for example due to capacity constraints or special-needs cases, introduces non-random sorting of job seekers to caseworkers, which in turn cause estimates of β to be biased. To address this concern, we follow Cederlöf et al. (2025) and implement a two-stage least squares (2SLS) strategy using the referral behavior of the predicted caseworker as an instrument for that of the assigned (actual) caseworker. Specifically, we compute a leave-out-mean for the predicted caseworker $\hat{j}$ based on job seekers born on the dates assigned to caseworker $\hat{j}$—analogous to the construction of the leave-out-mean among job seekers assigned to caseworker j (see equation (8)). Because day of birth is orthogonal to job seekers' observed and unobserved characteristics, predicted caseworkers are as good as randomly assigned a comparable pool of clients. Consequently, variation in referral behavior across predicted caseworkers reflects differences in gender (non-)stereotypical referral tendencies rather than differences in client composition, and since date-of-birth is used when assigning job seekers to caseworkers, the referral behavior of the predicted caseworker is predictive of that of the actual caseworker. We validate both of these conjectures below.

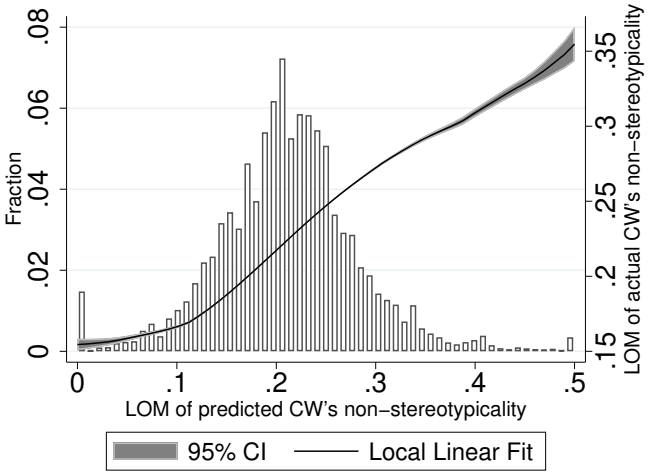
Assessing the instrument

First stage:

As noted earlier, the share of spells where the predicted caseworker corresponds to that of the actually assigned caseworker is 39%. This suggests that also referrals provided to job seekers assigned by their date-of-birth to a predicted caseworker should be informative about actual caseworker referral behavior. Figure 3 shows identifying variation in our instrument. The instrument is centered at 0.21 with a standard deviation of 0.08, but there is substantial variation across the distribution with predicted caseworkers in the 10^{th} and 90^{th} percentile sending 12% and 30% non-stereotypical referrals, respectively. Figure 3 also shows the assigned caseworker's non-stereotypical referral behavior against that of the predicted caseworker, providing a flexible analog of the first-stage relationship by plotting estimates from a local linear regression. The actual caseworker's leave-out mean increases roughly monotonically with the behavior of predicted caseworkers.

Appendix table A3 reports first-stage regression estimates where actual caseworker's leave-out-mean of non-stereotypical referral is regressed on predicted caseworker's non-stereotypical leave-out-mean. Column (1) suggests having a date-of-birth that renders a predicted caseworker one standard deviation up in

Figure 3. First stage graph using non-standardized leave-out mean



Notes: The leave-one-out-mean (LOM) is calculated by caseworker, office, and year. The interacted office and year fixed effects are used. Standard errors are clustered at office level. The sample restrictions used here are: (i) total number of job seekers per office per year>500, (ii) total number of caseworkers per office per year>5, and (iii) total number of referrals sent out by caseworkers during the sample period>50, and (iv) F-stat>200 offices are used. Gender split is based on 40-60 rules. Asterisks indicate significance at the * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ levels.

the non-stereotypical distribution, corresponds to an 0.12 standard deviation upward move in the assigned caseworker distribution. This estimate is highly significant and as an F-statistic well above conventional levels indicating weak instruments. In subsequent analysis, we investigate whether getting a non-stereotypical caseworker varies by the gender of the job seeker. We therefore report in column (2) the first-stage relationship separately for men and women, which, reassuringly, lines up well with the overall average effect.

Independence:

Because the instrument is constructed as a leave-out mean and assigned based on job seekers' date of birth, it should be independent of job seeker characteristics. To illustrate this, we correlate job seekers' predicted earnings with the (leave-out) referral behavior of both the actual and the predicted caseworker. Figure 4a shows that the referral behavior of the actual caseworker is strongly correlated with the type of job seekers they are assigned, suggesting non-random sorting of job seekers to caseworkers even within date-of-birth offices, for example through selective exemptions. However, as shown in figure 4b, this correlation disappears when predicted earnings are correlated with the referral behavior of the predicted caseworker.

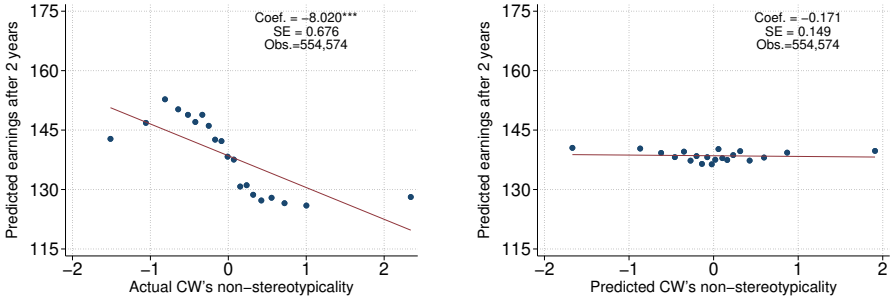
Appendix table A4 provides additional evidence on instrument independence by regressing the endogenous variable and the instrument (columns 1 and 2, respectively) on a set of job seeker characteristics. The referral behavior of the

actual caseworker is strongly correlated with nearly all characteristics.[7] For example, job seekers assigned to caseworkers one standard deviation higher in the non-stereotypical referral distribution are on average 2.3 years younger. In contrast, when regressing the instrument on the same characteristics, the coefficients are substantially smaller and jointly insignificant ($p = 0.360$). The only characteristic significantly correlated with the instrument at conventional levels is age. However, this correlation declines from 2.3 years to less than one month, suggesting that the remaining imbalance is economically negligible. Appendix table A5 shows that our results are unchanged when controlling for worker characteristics.

Taken together, these results support the independence of our instrument. While the referral behavior of the actual caseworker is strongly correlated with job seeker characteristics, the leave-out predicted caseworker behavior is not. Once assigned through the date-of-birth rule, the correlations with observed characteristics largely disappear and are jointly insignificant, suggesting that the instrument provides plausibly exogenous variation in exposure to non-stereotypical caseworker referral behavior.

Figure 4. Balancing figure w.r.t. predicted earnings after 2 years

(a) Using actual caseworker *(b)* Using predicted caseworker



Notes: Predicted earnings at $t+2$ versus actual (a) and predicted (b) caseworker non-stereotypicalness. The figure plots predicted earnings at $t+2$ against the non-stereotypicalness of (a) the actual caseworker, using data from all offices, and (b) the predicted caseworker, using data from the date-of-birth offices. The sample includes inflows of job seekers during 2007–2015. Points represent averages within equal-sized bins and are overlaid with a linear fit; point estimates and standard errors are reported in each panel. Predicted earnings at $t+2$ are constructed as fitted values from a regression of earnings at $t+2$ on age, education, marital status, children, employment one year prior to the spell, participation in any program during the previous four years, and the number of unemployment spells during the previous five years. Education is captured by indicators for pre-high school <9 years, pre-high school 9 years, high school 2 years, high school 3 years, university/college 3 years, and university/college >3 years or PhD.

Monotonicity:

Our IV approach requires monotonicity: assignment to a predicted caseworker with less stereotypical behavior should increase the likelihood of being assigned

[7] In table A6 in the appendix, we provide additional balancing tests interacting characteristics with job seeker gender.

an actual caseworker with less stereotypical behavior relative to assignment to a more stereotypical predicted caseworker.

Following Chyn et al. (2025), we test for average monotonicity by examining the first-stage relationship across different sub groups of the population, making sure the sign of the point estimate for one group (e.g, male job seekers) goes in the same direction as its complement (female job seekers).[8] We report these results in table A7 in the appendix, showing that across all subgroups, the first-stage coefficients are positive, statistically significant, and of similar magnitude. We interpret this as evidence providing strong support for the average monotonicity assumption.

Exclusion restriction:
The exclusion restriction in our setting implies that predicted caseworkers' referral behavior should not be correlated with other caseworker traits which in itself directly affects job seekers' labor market outcomes. For example, if non-stereotypical (predicted) caseworkers also are more prone to meet with their job seekers or have an overall higher propensity of providing them with referrals, our IV-estimates would not only reflect the effect of a caseworkers' non-stereotypical referral behavior but also other correlated traits.

While this assumption cannot be tested directly, we investigate the robustness of our results by measuring other caseworker traits including tenure, meeting propensity, assignment to active labor market programs, overall referral propensity, and preferences for referring job seekers to tight and broader occupations. We measure these traits in the same way as our main dependent variable and include them in our baseline regression, instrumented with the respective trait of the predicted caseworker. Table 6 shows that the estimates of caseworkers' non-stereotypical behavior is largely unaffected by the inclusion of other types of caseworker traits. We interpret this as providing some support in favor of the exclusion restriction.

Effects of a gender non-stereotypical caseworkers

Average effects

Table 3 shows that non-stereotypical caseworkers boost job-finding. The table presents IV estimates, where non-stereotypicalness is standardized to have a mean of zero and a standard deviation of one. The interpretation of the estimate in column (1) therefore is that a one standard deviation more non-stereotypical caseworker increases the probability of exiting unemployment

[8] Frandsen (2025) propose a test for the stricter assumption of *pairwise* monotonicity. However, for our purposes of identifying a LATE average monotonicity is sufficient.

within 180 days by 5.7 percentage points, or 10.4%. We regard this as an economically meaningful effect.

Table 3. *Effects of a non-stereotypical caseworkers on job seekers' outcomes*

	(1) Exit unemployment < 180 days	(2) Earnings 1 year spell start	(3) Earnings 2 years spell start	(4) Earnings 3 years spell start	(5) Earnings 4 years spell start	(6) Earnings 5 years after spell start
Non-stereotypical caseworker (standardized)	0.057*** (0.011)	6.519*** (1.727)	4.776*** (1.747)	3.313* (1.727)	3.571* (1.747)	4.260** (1.747)
Mean	0.547	115.061	138.155	154.350	165.341	169.141
# observations	554,574	554,574	554,574	554,574	554,574	554,574

Notes: IV estimates based on the regression model in 9 and the sample described in section 4. Outcomes are an indicator for exit from unemployment within 180 days and earnings measured one to five years after entry into unemployment. The explanatory variable is standardized (mean zero, standard deviation one) non-stereotypical caseworker behavior, instrumented using the non-stereotypical behavior of the predicted caseworker (both are leave-one-out measures). The models also include office×year fixed effects. Standard errors, clustered at the office level, are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The other columns of table 3 report effects on labor earnings. We have data on yearly earnings, and column (2) shows effects on yearly earnings in the calendar year after entry to unemployment and columns (3)–(6) effects for subsequent calendar years. The earnings effect of a more non-stereotypical caseworkers is positive and statistically significant across all horizons, but declines over time. For instance, a one standard deviation more non-stereotypical caseworker increases annual earnings by approximately 6,500 SEK (5.7%) one year after entry to unemployment and 4,200 SEK (2.5%) after five years. The persistence of the effects shows that the benefits of a non-stereotypical caseworkers extend beyond short-run reemployment and seem to translate into sustained improvements in employment outcomes.

We conclude that non-stereotypical caseworkers improve labor market outcomes. That is, pushing job seekers toward non-traditional occupations does not hurt them; it even helps them in the labor market. This suggests that policies aimed at breaking gender occupational segregation may not only create more gender-neutral labor markets but also improve labor market outcomes, at least for unemployed job seekers. We estimate the reduced-form effect of being assigned to a non-stereotypical caseworker, so we can only speculate about the exact underlying mechanisms. One explanation is that non-stereotypical caseworkers expand the opportunity set faced by job seekers by broadening the types of occupations on which they focus their search. By not considering non-traditional occupations, job seekers may forgo job opportunities, and it seems like non-stereotypical caseworkers remedy some of these gendered job-search constraints.

To shed some on these mechanisms, table 4 explores heterogeneity by prior occupational background, distinguishing between job seekers whose previous

occupation before becoming unemployed was gender-stereotypical (e.g., males in male-dominated occupations) and those whose prior occupation was non-stereotypical (e.g., males in female-dominated occupations). We do this by interacting our non-stereotypical caseworker variable with an indicator for a previous stereotypical occupation, and instrument these variables with the corresponding variables for the predicted caseworker.[9]

Table 4. *Heterogeneous effects of a non-stereotypical by previous occupation*

	(1) Unemployment Duration $<$ 180 days	(2) Earnings 1 year after spell start	(3) Earnings 2 years after spell start	(4) Earnings 3 years after spell start	(5) Earnings 4 years after spell start	(6) Earnings 5 years after spell start
Previous occupation stereotypical $\times$	0.046***	7.395**	7.145**	8.860**	8.328***	3.315
Non-stereotypical caseworker	(0.014)	(3.243)	(3.257)	(3.485)	(3.041)	(3.419)
Mean	0.542	138.605	160.655	175.361	184.205	186.264
Previous occupation not stereotypical $\times$	0.051***	8.383**	3.641	1.112	1.966	5.582
Non-stereotypical caseworker	(0.017)	(3.718)	(4.258)	(4.210)	(4.716)	(4.526)
Mean	0.549	146.488	172.375	189.844	199.262	197.128
# observations	316,939	316,939	316,939	316,939	316,939	316,939

Notes: The table reports estimates from regressions at the unemployment spell level. Outcomes include log unemployment duration, an indicator for unemployment duration shorter than 180 days, and earnings measured one to five years after spell start. The outcomes are regressed separately on non-stereotypical caseworker behavior (measured by the standardized leave-one-out mean of non-stereotypical referrals by the actual caseworker) and its interaction with indicators for whether the job seeker's previous occupation is stereotypical or not. These leave-one-out measures are instrumented by the non-stereotypical behavior of the predicted caseworker's predicted clients and are standardized to have mean zero and standard deviation one. The sample is restricted to job seekers with prior work experience. Additional sample restrictions are: (i) total number of job seekers per office per year $>$ 500, (ii) total number of caseworkers per office per year $>$ 5, (iii) total number of referrals sent by caseworkers during the sample period $>$ 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics $>$ 200 are used. Gender split is based on 40–60 rules. Interacted office–year–previous-occupation-stereotypical fixed effects are included. Standard errors in parentheses are clustered at the office level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The results from this exercise shows that job seekers with a stereotypical occupational background benefit more from a non-stereotypical caseworkers. The earnings gain after two years from a one-standard-deviation increase in non-stereotypicalness is 7,145 SEK and is significant for those coming from stereotypical occupation, and smaller and insignificant for those who previously worked in a non-stereotypical occupation. Assuming that job seekers from stereotypical occupations are more gender-constrained in their job search, these results support the interpretation that non-stereotypical caseworkers remedy some inefficient gendered job-search constraints. This is a likely explanation to the effects on job finding and labor earnings.

[9]Here, we get a slightly smaller sample as previous occupation is not observed for everyone. The reason is that the occupational data is based on annual survey data, available for roughly half of the sample, but we boost the sample by going back to the first observed occupation in the last five years.

Differences between male and female job seekers

Table 5 reveals pronounced gender heterogeneity in the effects of non-stereotypical caseworkers, with male job seekers benefiting more than female job seekers. In the table we interact caseworker non-stereotypicalness with indicators for male and female job seekers, and instrument these two endogenous variables with corresponding variables for the predicted caseworker.

For men, a one standard deviation more non-stereotypical caseworker increases the probability of exiting unemployment within 180 days by 7.6 percentage points, or 13.7% (column 1). The corresponding effect for female job seekers is more modest: 3.8 percentage points, or 7.1%. The difference between genders become even more pronounced when examining labor earnings in the more medium run. For men, labor earnings increase by around 9.6% after one year. This gain is statistically significant and remains positive and significant over the following five years. For women, there is no significant effect on labor earnings in the short or medium run.

Table 5. *Heterogeneous effects of a non-stereotypical caseworker for male and female job seekers*

	(1) Unemployment Duration < 180 days	(2) Earnings 1 year after spell start	(3) Earnings 2 years after spell start	(4) Earnings 3 years after spell start	(5) Earnings 4 years after spell start	(6) Earnings 5 years after spell start
Male × Non-stereotypical caseworker	0.076***	11.775***	6.866**	5.999**	8.888***	8.682***
	(0.015)	(2.780)	(2.681)	(2.810)	(3.107)	(3.116)
Mean	0.556	123.024	149.736	168.324	180.268	183.465
Female × Non-stereotypical caseworker	0.038***	0.974	2.435	0.473	-1.760	-0.228
	(0.011)	(2.040)	(2.320)	(2.565)	(2.458)	(2.707)
Mean	0.535	105.579	124.367	137.712	147.568	152.087
# observations	554,574	554,574	554,574	554,574	554,574	554,574

Notes: The table reports estimates from regressions at the unemployment spell level. Outcomes include log unemployment duration, an indicator for unemployment duration shorter than 180 days, and earnings measured one to five years after spell start. The outcomes are regressed separately on non-stereotypical caseworker behavior (measured by the standardized leave-one-out mean of non-stereotypical referrals by the actual caseworker) and its interaction with male and female indicators. These leave-one-out measures are instrumented by the non-stereotypical behavior of the predicted caseworker's predicted clients and are standardized to have mean zero and standard deviation one. The sample consists of job seekers who have received at least one referral and job seekers without receiving any referral. Additional sample restrictions are: (i) total number of job seekers per office per year > 500, (ii) total number of caseworkers per office per year > 5, (iii) total number of referrals sent by caseworkers during the sample period > 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics > 200 are used. Gender split is based on 40–60 rules. Interacted office–year–gender fixed effects are included. Standard errors in parentheses are clustered at the office level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

In sum, male job seekers experience the largest employment gains from being assigned to a non-stereotypical caseworker. It seems more important to relax binding occupational constraints and expand the set of viable job opportunities for male job seekers than for female job seekers. Thus, promoting more gender-balanced occupational choices and opening doors for women in male-dominated occupations may also help men, if it simultaneously opens

doors for them in female-dominated occupations. Thus, less gender segregation in the labor market has the potential to benefit both genders.

Non-stereotypicalness or other caseworker traits?

Caseworkers may differ along other dimensions that may be correlated with non-stereotypicalness. This holds both for caseworker characteristics (e.g., caseworkers with longer tenure may be more/less non-stereotypical) and actions that are not directly related to gender dimensions (e.g., number of meetings and assignment to labor market programs). To test for this, we have constructed detailed information on basically all actions taken by the caseworkers, and find that adjusting for these other factors cannot explain the observed differences between non-stereotypical and stereotypical caseworkers.

These robustness checks are presented in table 6 , starting with our baseline results in column (1), and then, columns (2)-(6), sequentially adds various other caseworker variables.[10] In terms of caseworker characteristics, we adjust for caseworker tenure, capturing experience and seniority. In terms of other caseworker strategies, we adjust for how frequent the caseworker meet with their job seekers and their propensity to assign job seekers to active labor market programs (ALMPs). Both, may be relevant if non-stereotypical caseworkers for some reason differ from other caseworkers in these two dimensions. Since, we use variation in the propensity to send non-stereotypical vacancy referrals, we also adjust for differences in the number of vacancy referrals that each caseworker send. Except for tenure, these caseworker variables are measured using the observed behavior of the caseworkers (e.g., how many meetings they have). As for non-stereotypicalness they are measured as leave-one-out means, and we instrument these additional caseworker variables using the corresponding variables for the predicted caseworker.

Table 6 clearly shows that the non-stereotypical caseworker estimates remain stable when adding these other caseworker variables. We conclude that our results indeed capture differences between stereotypical and non-stereotypical caseworkers, and not differences in some other correlated caseworker characteristics or strategies.

Non-traditional jobs, broader search and/or labor marker tightness?

Besides alleviating gendered constraints on search, promoting job search in non-stereotypical occupations implicitly also implies promoting search in other types of occupations. Non-stereotypical behavior may be linked to broader job search behavior. Applying to non-stereotypical jobs may simply reflect a wider

[10] Here, we focus on the male and female job seeker heterogeneity using earnings after two years as outcome table A8 report results for the main effect, and we reach similar conclusions using alternative outcomes.

search overall, and broader search—regardless of whether it targets stereotypical or non-stereotypical occupations—may itself be important (see, e.g., Belot et al., 2019). Additionally, applying to non-stereotypical jobs may also imply that you focus on more tight (slack) occupations with more (less) available jobs. For instance, the positive effects of non-stereotypical caseworkers for men may reflect that they are directed toward tighter occupations, as female-dominated occupations tend to be tighter than many male-dominated occupations, such as those in manufacturing. These three mechanisms–less gender-constrained search, broader search, and a focus on tighter occupations–are all part of the reduced-form effect of non-stereotypical caseworkers. However, their relative importance has implications for policy. For instance, should policy aim to promote non-stereotypical jobs, or is it sufficient to encourage broader search?

Table 6. *Adjusting for other caseworker traits and different mechanisms*

	Dependent variable: Earnings after 2 years						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Main	Other caseworker variables				Mechanisms	
		Tenure	Meeting	ALMP	Referral	Tightness	Breadth
Male × Non-stereotypical caseworker	6.866**	6.862**	6.849**	6.844**	6.850**	6.151**	4.778*
	(2.681)	(2.682)	(2.686)	(2.679)	(2.683)	(2.701)	(2.815)
Female × Non-stereotypical caseworker	2.435	2.430	2.395	2.404	2.379	1.528	-0.196
	(2.320)	(2.318)	(2.316)	(2.317)	(2.305)	(2.340)	(2.363)
# observations	554,574	554,574	554,574	554,574	554,574	554,574	554,574

Notes: The table reports robustness checks for the main specification using regressions at the unemployment spell level. The dependent variable is earnings two years after unemployment entry. Each column adds the indicated caseworker trait to the baseline specification separately. Non-stereotypical caseworker behavior is measured by the standardized leave-one-out mean of non-stereotypical referrals by the actual caseworker and is instrumented using the non-stereotypical behavior of the predicted caseworker's predicted clients. Standard errors in parentheses are clustered at the office level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Disentangling these mechanisms is challenging, but we attempt to do so using individual caseworker data on referrals to non-stereotypical occupations, broader sets of occupations, and tight versus slack occupations. The latter two are measured in a similar way as non-stereotypicalness. First, we characterize all referrals sent by caseworkers as either broad or narrow. A broad referral is defined as a referral to an occupation outside the job seeker's preferred set of occupations. Second, we use population data on vacancies and unemployment to measure labor market tightness for each occupation. This occupation-level measure is used to characterize the tightness of each referral. We then construct measures of broad and tight referral behavior for each caseworker based on the referrals they send in a similar way as for non-stereotypical referrals. As before, we use leave-one-out means and instrument these caseworker behaviors using the corresponding variables for the predicted caseworker. Given that we have sufficient variation in our measures of non-stereotypicalness, broader referrals, and occupational tightness, we can shed light on the relative importance of

these three mechanisms. However, since this places substantial demands on the variation in the data, the results should be interpreted with some caution.

The results from this exercise are presented in columns (6)-(7) of table 6, focusing on the male and female job seeker heterogeneity and earnings after 2 years as the main outcome. The effect for female job seekers was small and statistically insignificant even before controlling for breadth and tightness, and, as expected, it remained so after these controls were included. For male job seekers, the initial earnings estimate is 6,850 SEK. It drops slightly to 6,151 SEK when adjusting for tightness and further to 4,778 SEK when controlling for the caseworker's broader referral behavior. This pattern suggests that while non-stereotypicalness is correlated with broader and tighter referrals, it also has an independent effect beyond these channels. We conclude that promoting non-stereotypical job search help unemployed job seekers.

Effects on wages and occupations

Female job seekers

One motivation for the interest in gender occupational segregation is that shifts toward less segregated labor markets may reduce gender wage-gaps. We therefore study how a non-stereotypical caseworker affects subsequent wages and occupational choices. We begin by analyzing the type of occupations the job seekers end up in, examining effects on finding a job in a male- and female-dominated occupation (defined as above), respectively. Columns (3) of table 7 show that female job seekers are more likely to enter a male-dominated occupation if they get a more non-stereotypical caseworker, consistent with the idea that non-stereotypical caseworker broadens the set of viable jobs and encourages women to enter male-dominated occupations. The effect on a job in a female-dominated occupation is negative, as expected, but insignificant.

Furthermore, column (2) shows that non-stereotypical caseworkers lead to higher wages, presumably because it induces female job seekers to enter male-dominated occupations, which tend to be higher paid. However, both the wage and the male-dominated occupation effects are rather modest: a one-standard-deviation increase in caseworker non-stereotypicalness increases the probability of entering a male-dominated occupation by 1.8 percentage points (14.52%) and increase wages in the first job after unemployment by around 0.8 percent. It demonstrates that breaking gender segregation is challenging, but also that non-stereotypical caseworkers may be part of the solution.

Before turning to the results for male job seekers, note that we use survey data on wages and occupations. The benefit is that we have actual wages and not wages derived from earnings and hours worked. However, it also means

Table 7. *Effects on wages and future occupations for male and female job seekers*

	(1) Pr(wage observed)	(2) Log first wage	(3) New occupation male-dominated	(4) New occupation female-dominated
Male × Non-stereotypical caseworker	-0.006	0.005	0.010	-0.019
	(0.009)	(0.008)	(0.021)	(0.019)
Mean	0.293	10.070	0.543	0.358
Female × Non-stereotypical caseworker	-0.018*	0.008*	0.018*	-0.008
	(0.010)	(0.005)	(0.010)	(0.014)
Mean	0.429	10.024	0.124	0.795
# observations	554,574	196,183	196,126	196,126

Notes: The table reports estimates from regressions at the unemployment spell level. The outcomes include whether wages are observed, log first wage, and indicators for whether the new occupation is male- or female-dominated. Outcomes are regressed separately on non-stereotypical caseworker behavior (measured by the standardized leave-one-out mean of non-stereotypical referrals by the actual caseworker) and its interaction with male and female indicators. These leave-one-out measures are instrumented by the non-stereotypical behavior of the predicted caseworker's predicted clients and are standardized to have mean zero and standard deviation one. The sample consists of job seekers who have received at least one referral and job seekers without receiving any referral. Additional sample restrictions are: (i) total number of job seekers per office per year > 500, (ii) total number of caseworkers per office per year > 5, (iii) total number of referrals sent by caseworkers during the sample period > 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics > 200 are used. Gender split is based on 40–60 rules. Interacted office–year–gender fixed effects are included. Standard errors in parentheses are clustered at the office level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

that we only have wage data for a subsample[11], explaining the lower sample size. Column 1 of table 7 shows that having non-stereotypical caseworkers decreases the probability that we observe the wage. The reason is that the wage survey covers the entire public sector but only a sample of private firms, so fewer observed wages are consistent with more jobs in male-dominated occupations, which are primarily located in the private sector. However, it still means that wage estimates should be interpreted with caution.

Male job seekers

For male job seekers, the effect on wages is theoretically more ambiguous. A non-stereotypical caseworker may encourage male job seekers to enter female-dominated occupations, which are typically lower paid, but may also broaden the set of viable job opportunities and allow them to find better jobs. These counteracting factors may explain why we see no impact on the wages (table 7, column 2). This also implies that the earnings effect observed above must be driven by higher employment (faster job finding and/or more persistent employment).[12]

[11] To mitigate this, we use the first observed wage within 5 years, but there is still not full coverage.

[12] We have also examined the probability that the wage is observed and for the male job seekers we find no differences between those with stereotypical and non-stereotypical caseworkers (table 7, column 1).

Interestingly, column (3)–(4) of table 7 reveals no significant effect for male job seekers on the probability of entering either male- or female-dominated occupations, despite the fact that male job seekers benefit from non-stereotypical caseworkers. Several things may contribute to this finding. First, as already mentioned, we have occupational data for a smaller sample (36%) than for earnings, and it may be that we do not capture all the occupational changes. Second, job search and re-employment after unemployment involve several dimensions, including the choice of occupation, workplace, and industry. Non-stereotypical caseworkers may therefore help job seekers broaden their perspectives in the labor market, for example by encouraging them to consider different types of workplaces, even if they do not transition into entirely new occupations. Third, note that we measure the reduced-form effect of the non-stereotypical caseworker trait, which captures all actions they take, including how they interact with male and female job seekers during meetings, and not only their referral behavior. This may also contribute to the positive effects of a non-stereotypical caseworker for male job seekers.

Conclusions

This paper studies whether public employment services can help mitigate occupational gender segregation by shaping job seekers' exposure to gender-atypical employment opportunities. Focusing on the Swedish Public Employment Service, we provide causal evidence that caseworkers' non-stereotypical gender norms, i.e., their willingness to deviate from gender-stereotypical occupational matches, has meaningful effects on job seekers' labor market outcomes. Exploiting quasi-random assignment of job seekers to caseworkers based on date-of-birth rules and an instrumental variables strategy, we isolate exogenous variation in exposure to gender non-stereotypical referral environments.

Our main findings show that job seekers assigned to more gender non-stereotypical caseworkers exit unemployment faster and experience higher post-unemployment earnings. A one standard deviation increase in non-stereotypical referral behavior raises the probability of reemployment within six months by 5.7 percentage points and generates persistent earnings gains lasting up to five years after the start of the unemployment spell. These results may point to improvements in match quality rather than mere increases in job-finding rates, suggesting that non-stereotypical referrals expand job seekers' opportunity sets and facilitate more productive matches.

The heterogeneity analyses shed light on who benefits most from breaking gender norms in job referrals. The gains are largest for job seekers whose prior labor market trajectories were more strongly shaped by gender norms, particularly men and individuals with previously stereotypical occupational backgrounds. For these groups, non-stereotypical referrals appear to relax binding occupational constraints and reduce congestion in narrowly defined job queues, leading to both faster reemployment and higher long-run earnings. In

contrast, job seekers with more diverse or already non-stereotypical work histories experience smaller marginal benefits, consistent with diminishing returns once occupational boundaries have already been crossed. We also find that women exhibit a different adjustment margin: exposure to non-stereotypical caseworkers increases transitions into male-dominated occupations and leads to modest but statistically significant gains in first wages. These transitions are accompanied by slightly slower job matching, plausibly reflecting higher search frictions as women move toward less traditional and on average better-paying occupational pathways.

Our findings contribute to several strands of the literature. First, they complement work on occupational segregation and the gender wage gap by highlighting the role of labor market intermediaries as a mechanism shaping occupational sorting, alongside worker preferences and employer behavior. While prior research has emphasized sorting across firms and occupations as key drivers of gender wage differentials (e.g., Card et al., 2016; Goldin, 2014; Bruns, 2019), our results show that public employment services can actively influence these sorting processes at the job-search stage. Second, we add to a growing literature documenting persistent gender differences in job search behavior and application patterns (e.g., Barbulescu and Bidwell, 2013; Fluchtmann et al., 2024) by demonstrating that institutional guidance can partially counteract these patterns. Third, our results relate to recent evidence showing that removing or relaxing gender cues in labor market institutions, such as vacancy postings or recruitment practices, can reduce gender segregation (Card et al., 2025).

From a policy perspective, the results underscore that addressing occupational segregation requires more than increasing labor force participation. Even in a country like Sweden, with relatively high gender equality and extensive labor market institutions, gender norms continue to shape job search behavior and matching outcomes. Our findings align with the World Bank's (2023) conclusion that policies are most effective when they address both supply-side and demand-side constraints. By influencing job seekers' perceptions of what constitutes a viable job, caseworkers can play a pivotal role in broadening occupational choices and improving labor market efficiency.

Overall, this paper shows that breaking gender norms in job referrals yields tangible economic returns. By shaping the content of job search rather than merely its intensity, public employment services can help reduce occupational segregation, improve matching efficiency, and promote more equitable labor market outcomes. These insights suggest that the potential gains from norm-challenging employment services may be even larger in settings with stronger occupational segregation and weaker institutional support than those studied here.

References

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Bamieh, O. and Ziegler, L. (2024). Wage transparency and gender sorting. *Economic Policy*, pages 1–26.

Barbulescu, R. and Bidwell, M. (2013). Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science*, 24(3):737–756.

Blau, F. D. (2025). Gender inequality in the labor market: Continuing progress? *ILR Review*, 78(2):275–303.

Blau, F. D., Brummund, P., and Liu, A. Y.-H. (2013). Trends in occupational segregation by gender 1970–2009: Adjusting for the impact of changes in the occupational coding system. *Demography*, 50(2):471–492.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, 55(3):789–865.

Bruns, B. (2019). Changes in workplace heterogeneity and how they widen the gender wage gap. *American Economic Journal: Applied Economics*, 11(2):74–113.

Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drott, D. (2024). How are gender norms perceived? Manuscript, November 7, 2024.

Card, D., Cardoso, A. R., and Kline, P. (2015). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women *. *The Quarterly Journal of Economics*, 131(2):633–686.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly journal of economics*, 131(2):633–686.

Card, D., Colella, F., and Lalive, R. (2025). Gender preferences in job vacancies and workplace gender diversity. *Review of Economic Studies*, 92(4):2437–2471.

Carranza, E., Das, S., and Kotikula, A. (2023). Gender-based employment segregation: Understanding causes and policy interventions.

Cederlöf, J., Söderström, M., and Vikström, J. (2025). The role of caseworkers: Job finding, job quality and determinants of value-added. *Journal of the European Economic Association*, page jvaf055.

Cejka, M. A. and Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, 25(4):413–423.

Chyn, E., Frandsen, B., and Leslie, E. (2025). Examiner and judge designs in economics: a practitioner's guide. *Journal of Economic Literature*, 63(2):401–439.

Cortés, P., Hwang, J., Pan, J., and Schönberg, U. (2026). Gender norms and the labor market. NBER Working Paper 34716, National Bureau of Economic Research.

Cortes, P. and Pan, J. (2018). Occupation and gender. In *The Oxford Handbook of Women and the Economy*. Oxford University Press.

Cortés, P., Pan, J., Pilossoph, L., Reuben, E., and Zafar, B. (2023). Gender differences in job search and the earnings gap: Evidence from the field and lab. *The Quarterly Journal of Economics*, 138(4):2069–2126.

Del Carpio, L. and Guadalupe, M. (2022). More women in tech? evidence from a field experiment addressing social identity. *Management Science*, 68(5):3196–3218.

Delfino, A. (2024). Breaking gender barriers: Experimental evidence on men in pink-collar jobs. *American Economic Review*, 114(6):1816–1853.

England, P., Levine, A., and Mishel, E. (2020). Progress toward gender equality in the united states has slowed or stalled. *Proceedings of the National Academy of Sciences*, 117(13):6990–6997.

Fluchtmann, J., Glenny, A. M., Harmon, N. A., and Maibom, J. (2024). The gender application gap: Do men and women apply for the same jobs? *American Economic Journal: Economic Policy*, 16(2):182–219.

Foster, T. B., Murray-Close, M., Landivar, L. C., DeWolf, M., et al. (2020). *An evaluation of the gender wage gap using linked survey and administrative data*. US Census Bureau, Center for Economic Studies Washington, DC.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American economic review*, 104(4):1091–1119.

Hegewisch, A. and Hartmann, H. (2014). Occupational segregation and the gender wage gap: A job half done. Working paper, Institute for Women's Policy Research.

Hegewisch, A., Liepmann, H., Hayes, J., and Hartmann, H. (2010). Separate and not equal? gender segregation in the labor market and the gender wage gap. *IWPR Briefing Paper*, 377(September):1–16.

Humlum, A., Munch, J., and Rasmussen, M. (2025). What works for the unemployed? evidence from quasi-random caseworker assignments. Technical report, National Bureau of Economic Research.

Lagerström, J. (2011). How important are caseworkers-and why? new evidence from swedish employment offices. Technical report, Working Paper.

Liepmann, H. and Hegewisch, A. (2025). Revisiting occupational segregation and the valuation of women's work. Ilo working paper, International Labour Organization.

Lundin, D. (2004). Vad styr arbetsförmedlarna? Technical report, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Olivetti, C., Pan, J., and Petrongolo, B. (2024). The evolution of gender in the labor market. In *Handbook of labor economics*, volume 5, pages 619–677. Elsevier.

Pan, J. (2015). Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics*, 33(2):365–408.

Schiprowski, A. (2020). The role of caseworkers in unemployment insurance: Evidence from unplanned absences. *Journal of Labor Economics*, 38(4):1189–1225.

Sorkin, I. (2017). The role of firms in gender earnings inequality: Evidence from the united states. *American Economic Review*, 107(5):384–387.

World Bank (2023). Gender-based employment segregation: Understanding causes and policy interventions. Jobs Working Paper 26, World Bank, Washington, DC. Draft date: July 11, 2023.

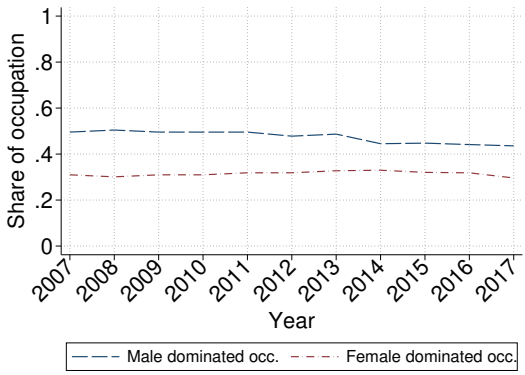
Appendix

Table A1. *Top 10 male and female dominated occupations*

	Wage survey data		Job-search data	
	Realized occupation (SSYK)	Female share	Preferred occupation (SSYK)	Female share
Panel A: Top 10 Female-Dominated Occupations				
1	Office secretaries and data entry clerks (411)	0.956	Midwives and nurses with special skills (223)	0.909
2	Preschool and recreational teachers (331)	0.910	Nurses (323)	0.899
3	Bookkeeping and accounting assistants (412)	0.892	Office secretaries and data entry clerks (411)	0.868
4	Nurses (323)	0.891	Hairdressers and personal service workers (514)	0.867
5	Midwives and nurses with special skills (223)	0.876	Preschool and recreational teachers (331)	0.841
6	Biomedical analysts (324)	0.872	Other office staff (419)	0.841
7	Physiotherapists and dental hygienists (322)	0.859	Accountants and administrative assistants (343)	0.838
8	Health and social care workers (513)	0.854	Customer information officers (422)	0.836
9	Psychologists and social workers (249)	0.823	Animal breeders and keepers (612)	0.833
10	Special education teachers (234)	0.811	Psychologists and social workers (249)	0.830
Panel B: Top 10 Male-Dominated Occupations				
1	Building and construction workers (712)	0.027	Building and construction workers (712)	0.014
2	Machine and engine repairers (723)	0.039	Building and construction laborers (931)	0.014
3	Casters and welders (721)	0.048	Machine and engine repairers (723)	0.023
4	Blacksmiths and tool makers (722)	0.071	Casters and welders (721)	0.034
5	Construction laborers (931)	0.082	Electrical and electronics repairers (724)	0.035
6	Military personnel (011)	0.086	Forestry workers (614)	0.037
7	Electrical and electronics repairers (724)	0.089	Fishermen and hunters (615)	0.038
8	Tyre fitters (834)	0.090	Building craftsmen (713)	0.043
9	Building craftsmen (713)	0.090	Operating machine operators (816)	0.056
10	Mining and quarrying workers (711)	0.091	Machine operators (833)	0.063

Notes: The sample consists of wage survey data and Swedish Public Employment Services (PES) administrative data for the years 2007–2013. Estimates are computed by pooling observations across years. Occupations follow the Swedish Standard Classification of Occupations 1996 (SSYK96). Missing data points with no pooled observations are omitted.

Figure A1. Share of male and female dominated occupation over time



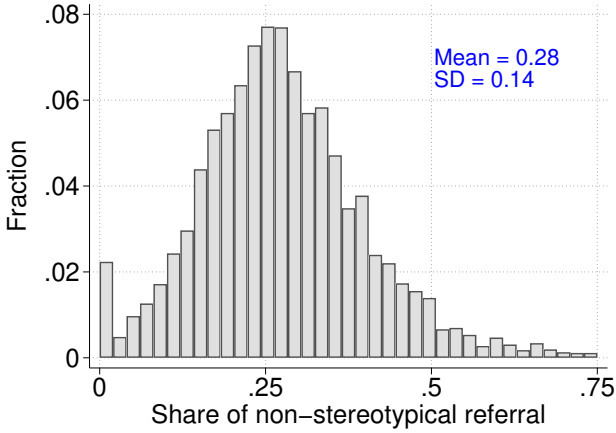
Notes: The figures show share of male and female dominated occupation in Sweden during 2007-2017. Gender share is calculated by dividing total male or female workers in a specific occupation by total workers in that occupation. Male (or female) dominated occupations are defined as having greater than sixty percent male (or female) workers in a specific occupation. The share of male and female dominated occupations is obtained by dividing the total male or female dominated occupations by total number of existing occupations each year. The figure is estimated based on the annual wage survey data in Sweden from 2007-2017, where use 3-digit occupational code from Statistics Sweden. The dashed blue and red lines represent the average male and female shares per occupation over time.

Table A2. *Descriptive statistics*

	Date-of-birth offices		All offices	
	Mean	SD	Mean	SD
Panel (a): Office characteristics				
Number of caseworkers	18.981	11.995	25.973	16.888
Number of unique job seekers	860.752	573.177	661.563	516.826
Number of job seeker spells	1072.677	712.033	972.922	724.265
# Obs. (office-year panel)	517		1,470	
Panel (b): Job seeker characteristics				
Male	0.544	0.498	0.550	0.497
Age	31.883	12.840	32.162	13.027
Pre-high school $<$ 9 years	0.014	0.116	0.015	0.120
Pre-high school 9 years	0.288	0.453	0.292	0.455
High school 2 years	0.173	0.378	0.181	0.385
High school 3 years	0.311	0.463	0.310	0.463
University college 3 years	0.098	0.298	0.094	0.292
University college $\geq$ 3 years / PhD	0.110	0.313	0.101	0.301
Married	0.169	0.374	0.171	0.376
Children	0.308	0.462	0.309	0.462
Employed 1 year before spell	0.560	0.496	0.555	0.497
Any program in last 4 years	0.034	0.181	0.038	0.191
Number of unemployment spells (5 yrs)	1.390	1.890	1.397	1.912
# Obs. (spell level panel)	554,574		1,430,195	
Panel (c): Caseworker characteristics				
Referral behavior:				
Total referrals*	63.283	90.293	51.966	84.620
Share of non-stereotypical referrals*	0.212	0.179	0.212	0.223
Meetings and workload:				
Number of meetings (annual)	209.010	224.258	152.058	204.573
Number of job seekers per year	57.955	61.779	41.818	55.566
Share of job seekers assigned to ALMP	0.250	0.201	0.271	0.256
Tenure:				
0–2 years	0.010	0.099	0.013	0.112
2–5 years	0.382	0.486	0.328	0.470
5–10 years	0.526	0.499	0.526	0.499
10+ years	0.083	0.275	0.133	0.339
# Obs. (caseworker-year panel)	9,569		34,200	

Notes: The table reports means and standard deviations for office-, job seeker-, and caseworker-level characteristics in date-of-birth offices and all offices. Job seeker characteristics are measured at the start of the unemployment spell. '*' denotes that the caseworker-year panel includes only caseworkers who give at least one referral (here, # Obs. are 9,567 and 28,788 for date-of-birth offices and all offices, respectively).

Figure A2. Share of non-stereotypical referrals sent by caseworkers



Notes: The above figure is computed using the total number of non-stereotypical occupational referrals sent out by caseworkers pooled across the entire sample period 2007-2017. The sample restrictions used here are: (i) the sample only consists of spells with at least one referral, (ii) total number of job seekers per office per year>500, (iii) total number of caseworkers per office per year>5, and (iv) total number of referrals sent out by caseworkers during the sample period>50, (v) immigrants are excluded from the sample, and (vi) the data are at the caseworker level (pooled across years).

Table A3. *First stage estimates for non-stereotypical behavior using Regular standardized instrument*

	Dependent variable: Actual non-stereotypical caseworker (standardized)	
	(1)	(2)
All jobseekers		
Predicted non-stereotypical caseworker (standardized)	0.120***	
	(0.011)	
Male and female jobseekers		
Male × Predicted non-stereotypical caseworker (standardized)		0.115***
		(0.012)
Female × Predicted non-stereotypical caseworker (standardized)		0.126***
		(0.012)
F-statistic	109.749	53.985
# observations	554,574	554,574

Notes: The leave-one-out mean (LOM) is calculated by caseworker, office, and year. The instrument is standardized with mean zero and standard deviation one. Interacted office and year fixed effects are used in column (1), and interacted office, year, and gender fixed effects are used in column (2). Standard errors are clustered at the office level. The sample restrictions used here are: (i) total number of job seekers per office per year $>$ 500, (ii) total number of caseworkers per office per year $>$ 5, (iii) total number of referrals sent out by caseworkers during the sample period $>$ 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics $>$ 200 are used. Gender split is based on 40–60 rules. The data are at the spell level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A4. *Balancing regression using regular LOM*

	Dependent variables	
	(1) Actual non-stereotypical caseworker (standardized)	(2) Predicted non-stereotypical caseworker (standardized)
Job-seeker characteristics:		
Male	-0.018***	0.001
	(0.004)	(0.001)
Age	-2.278***	-0.097***
	(0.231)	(0.032)
Pre-high school $<$ 9 yrs	-0.001	-0.000
	(0.000)	(0.000)
Pre-high school 9 yrs	0.051***	0.002*
	(0.005)	(0.001)
High school 2 yrs	-0.026***	-0.000
	(0.003)	(0.001)
High school 3 yrs	0.007***	0.000
	(0.002)	(0.001)
Univ. college 3 yrs	-0.010***	-0.001
	(0.001)	(0.001)
Univ. college $\geq$ 3 yrs / PhD	-0.022***	-0.001
	(0.003)	(0.001)
Married	-0.030***	-0.002*
	(0.004)	(0.001)
Child	-0.008***	-0.001
	(0.002)	(0.001)
Employed 1 year before spell	-0.064***	-0.001
	(0.005)	(0.001)
Any program in last 4 years	0.001***	-0.000*
	(0.000)	(0.000)
Number of unemployment spells (5 yrs)	-0.114***	-0.000
	(0.013)	(0.004)
F-statistic	26.701	1.102
p-value	0.000	0.360
Observations	554,574	554,574

Notes: The table reports separate OLS regressions of each job seeker characteristic on actual and predicted caseworker stereotypical behavior, measured as standardized leave-one-out means of non-stereotypical referrals. The sample consists of unique unemployment spells in Sweden from 2007–2015 that received job referrals (referral type 2) during 2007–2017. Interacted office and year fixed effects are included. Standard errors are clustered at the office level. Sample restrictions are: (i) total number of job seekers per office per year $>$ 500, (ii) total number of caseworkers per office per year $>$ 5, (iii) total number of referrals sent by caseworkers during the sample period $>$ 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics $>$ 200 are used. Gender split is based on 40–60 rules. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A5. *Balancing regression using regular LOM (controlling for job-seekers characteristics)*

	Dependent variables	
	(1) Actual non-stereotypical caseworker (standardized)	(2) Predicted non-stereotypical caseworker (standardized)
Job-seeker characteristics:		
Male	-0.064***	0.001
	(0.009)	(0.002)
Age	-0.008***	-0.000***
	(0.001)	(0.000)
Pre-high school $<$ 9 yrs	-0.019	-0.017
	(0.036)	(0.018)
Pre-high school 9 yrs	-0.074**	-0.013
	(0.035)	(0.016)
High school 2 yrs	-0.138***	-0.014
	(0.033)	(0.017)
High school 3 yrs	-0.131***	-0.017
	(0.035)	(0.016)
Univ. college 3 yrs	-0.175***	-0.018
	(0.034)	(0.018)
Univ. college $\geq$ 3 yrs / PhD	-0.236***	-0.019
	(0.037)	(0.017)
Married	0.018***	-0.000
	(0.005)	(0.003)
Child	-0.045***	-0.001
	(0.005)	(0.002)
Employed 1 year before spell	-0.081***	0.002
	(0.006)	(0.003)
Any program in last 4 years	0.012*	-0.008
	(0.007)	(0.005)
Number of unemployment spells (5 yrs)	-0.014***	0.000
	(0.001)	(0.001)
F-statistic	26.701	1.102
p-value	0.000	0.360
Observations	554,574	554,574

Notes: The table reports the results from joint OLS regressions of the actual and predicted caseworkers non-stereotypicalness (measured as standardized leave-one-out means of non-stereotypical referrals) on all job seeker characteristics. The sample consists of unique unemployment spells in Sweden from 2007–2015 that received job referrals (referral type 2) during 2007–2017. Interacted office and year fixed effects are included. Standard errors are clustered at the office level. Sample restrictions are: (i) total number of job seekers per office per year $>$ 500, (ii) total number of caseworkers per office per year $>$ 5, (iii) total number of referrals sent by caseworkers during the sample period $>$ 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics $>$ 200 are used. Gender split is based on 40–60 rules. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A6. *Balancing regression using regular LOM interacted with gender*

	Dependent variable: *Predicted caseworker stereotypical behavior*	
	Males (1)	Females (2)
Job-seeker characteristics:		
Age	-0.098**	-0.097***
	(0.043)	(0.036)
Pre-high school $<$ 9 yrs	-0.000	-0.000
	(0.000)	(0.000)
Pre-high school 9 yrs	0.003**	0.000
	(0.001)	(0.001)
High school 2 yrs	-0.001	0.000
	(0.001)	(0.001)
High school 3 yrs	-0.000	0.000
	(0.001)	(0.001)
Univ. college 3 yrs	-0.001**	0.000
	(0.001)	(0.001)
Univ. college $\geq$ 3 yrs / PhD	-0.000	-0.001
	(0.001)	(0.001)
Married	-0.002*	0.001
	(0.001)	(0.001)
Child	-0.001	-0.000
	(0.001)	(0.001)
Employed 1 year before spell	-0.001	-0.001
	(0.001)	(0.002)
Any program in last 4 years	-0.000	-0.000
	(0.000)	(0.000)
Number of unemployment spells (5 yrs)	0.005	-0.008*
	(0.007)	(0.004)
F-statistic	1.378	
p-value	0.125	
Observations	554,574	

Notes: The table reports separate OLS regressions of each job seeker characteristic on actual/predicted caseworker stereotypical behavior, measured as the standardized leave-one-out mean of non-stereotypical referrals. The sample consists of unique unemployment spells in Sweden from 2007–2015 that received job referrals (referral type 2) during 2007–2017. Interacted office–year–gender fixed effects are included. Standard errors are clustered at the office level. Sample restrictions are: (i) total number of job seekers per office per year $>$ 500, (ii) total number of caseworkers per office per year $>$ 5, (iii) total number of referrals sent by caseworkers during the sample period $>$ 50, (iv) immigrants are excluded from the sample, and (v) offices with F-statistics $>$ 200 are used. Gender split is based on 40–60 rules. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A7. *Monotonicity Test: First-stage estimates using complementary-subgroup instrument*

	Dependent variable: Actual non-stereotypical caseworker (standardized)				
	Male	Female	Married	Not married	Age < 40
Predicted non-stereotypical caseworker (standardized)	0.132*** (0.014)	0.156*** (0.016)	0.118*** (0.013)	0.193*** (0.017)	0.174*** (0.022)
Observations	237,028	219,856	87,989	236,891	217,498
F-statistic	91.140	90.408	80.650	128.385	62.264
	Age ≥ 40	Have child	No child	Program	No program
Predicted non-stereotypical caseworker (standardized)	0.127*** (0.014)	0.131*** (0.014)	0.154*** (0.013)	0.128*** (0.022)	0.202*** (0.026)
Observations	126,524	160,909	307,430	17,840	125,043
F-statistic	84.662	93.412	132.716	33.409	62.511

Notes: This table reports first-stage estimates from regressions of actual caseworker non-stereotypical behavior on predicted caseworker non-stereotypical behavior, where the instrument is constructed within a complementary subgroup. Specifically, the leave-one-out mean (LOM) of predicted caseworker non-stereotypicalness is calculated by caseworker, office, and year using job seekers belonging to the complementary subgroup, and then used as an instrument in the subgroup indicated by the column heading. The instrument is standardized to have mean zero and unit variance. All specifications include interacted office and year fixed effects. Standard errors are clustered at the office level. The sample is restricted to offices with more than 500 job seekers per year, more than 5 caseworkers per year, and caseworkers with more than 50 referrals during the sample period. Immigrants are excluded, and the analysis is conducted at the spell level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A8. *Robustness analyses adjusting for other caseworker traits/strategies*

	Dependent variable: Earnings after 2 years						
	(1) Main	(2)	(3)	(4)	(5)	(6)	(7)
		Other caseworker variables				Mechanisms	
		Tenure	Meeting	ALMP	Referral	Tightness	Breadth
Non-stereotypical caseworker	4.939*** (1.774)	4.934*** (1.774)	4.922*** (1.774)	4.924*** (1.776)	4.915*** (1.776)	4.140** (1.806)	2.543 (1.892)
# observations	554,574	554,574	554,574	554,574	554,574	554,574	554,574

Notes: The table reports robustness checks for the main specification using regressions at the unemployment spell level. The dependent variable is earnings two years after unemployment entry. Each column adds the indicated caseworker trait to the baseline specification separately. Non-stereotypical caseworker behavior is measured by the standardized leave-one-out mean of non-stereotypical referrals by the actual caseworker and is instrumented using the non-stereotypical behavior of the predicted caseworkers' predicted clients. Standard errors in parentheses are clustered at the office level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Financing Secondary Education: Closing Gaps or Boosting Grades?[1]

[1]I would like to thank Johan Vikström, Peter Fredriksson, Pierre Deschamps, Olof Rosenqvist, Martin Mattsson, José Montalbán Castilla, Md. Mojibur Rahman, Shahriar Jamil, Md. Zakir Hossen, Mahbub Zaman Khan, Mohd. Tari1ul Islam, Monjur-E-Morshed, Oskar Nordström Skans, Lena Hensvik, Georg Graetz, Jonas Cederlöf, and Eva Mörk for their insightful comments and suggestions. This work has further benefited from discussions with members of the Uppsala Labor Group.

Introduction

Primary and secondary education are arguably crucial, as they lay the foundation for higher education and skill acquisition. Governments worldwide recognize this importance and invest heavily in school education to improve access, quality, and equity (Psacharopoulos, 1994; Hanushek, 2006). Despite this effort, the effectiveness of public spending in education has been debated extensively in the literature (e.g., Dustmann et al., 2003; Hoxby, 2000; Marshall, 2009). Previous research underscores that the returns to educational investment are profoundly shaped by context.[2] Identical interventions (e.g., infrastructure, teacher quality, instructional resources, administrative capacity etc.) can generate markedly different outcomes depending on prevailing socioeconomic conditions, institutional arrangements, and political environments. These variations often reflect differences in governance quality, bureaucratic capacity, and accountability structures. Moreover, the distinction between capital and recurrent expenditures, and between investments in physical infrastructure versus those aimed at strengthening pedagogical quality may also play a role.

This paper examines the impact of public investment in the secondary education sector on school outcomes in Bangladesh, where resources were allocated under a system marked by inadequate monitoring, minimal restrictions, and a self-sustaining, long-term commitment. The analysis focuses on the Monthly Pay Order (MPO) program, one of the most significant and enduring government interventions in Bangladesh's education sector, particularly for non-government educational institutions in rural areas under which the government funds the basic payroll expenses of the teachers.[3] Currently, the government finances at least two-thirds of basic payroll expenditures in over 90 percent of MPO-recipient institutions and fully covers the basic costs in more than half of them. By securing financial sustainability through this arrangement, the program aims to strengthen teaching quality and promote greater stability within the education workforce.

To identify the causal impact of the payroll subsidy program, I leverage the program's clear eligibility thresholds based on schools' prior exam passing rates and number of students appeared in the public exam. Schools just above these cutoffs received the subsidy, while those just below did not, creating a natural comparison between otherwise similar schools. By focusing on this narrow margin, I can attribute any discontinuous changes in outcomes to the

[2]See Baron (2022); Martorell et al. (2016); Jackson and Mackevicius (2024); Lafortune and Schönholzer (2022) for details.

[3]Introduced in the 1980s, the MPO scheme was designed to provide financial security for teachers in non-government institutions, thereby sustaining the majority of the country's schools. By guaranteeing government-subsidized salaries, it aims to address teachers' livelihood needs and ensured continuity in academic activities. Beyond financial relief, the scheme also aims to promote transparency and merit-based recruitment through the Non-Government Teachers' Registration and Certification Authority (NTRCA), expand equitable promotion opportunities and professional recognition, and introduce job security measures.

program itself. This regression discontinuity framework thus offers a credible strategy for isolating the program's effects from other confounding factors. The design satisfies standard validity checks: first stage (showing deterministic jump in treatment probability at the cutoff), the running variable shows no evidence of manipulation[4], and baseline covariates are smooth across the cutoff. To estimate the local average treatment effects of the reform on major school-level educational outcomes, I use a newly compiled panel of school-level administrative data covering the period 2012–2023. To explore heterogeneity in treatment effects, shed light on potential mechanisms, and situate the findings in their broader context, I complement the baseline analysis with regional poverty headcount ratios from a nationally representative household survey to capture regional socioeconomic variation, education-sector corruption indicators, and event-level measures of MPO teacher dissatisfaction. Taken together, these data provide a richer basis for understanding both the reform's effects and the institutional environment in which they unfold.

Despite being a large and costly program, I find no significant effects of the MPO reform on students' academic performance. The major performance indicators (i.e., passing rate and share of distinctions in secondary school completion exam) show no statistically significant improvements among MPO-recipient institutions. These findings suggest that the payroll subsidy program did not generate measurable gains in student achievement. Nevertheless, the estimates indicate a positive post-reform effect on the girls' secondary school participation rate in treated schools, of 0.21 percentage points under the first running variable and 0.15 percentage points under the second[5]. While this pattern is consistent with a modest expansion in girls' access to secondary schooling, the evidence is not strong enough to support a more definitive conclusion. I also examine heterogeneity by regional poverty and educational corruption, but find no meaningful differences in outcomes across these spatial dimensions. As a robustness check, I employ a multi-score regression discontinuity design following Cattaneo et al. (2020b) to assess the sensitivity of the results (appendix B). The results are broadly in line with the parametric sharp RDD estimates, which similarly fail to detect statistically meaningful treatment effects on performance outcome but experience modest gain in girl's secondary school participation in some cases[6].

The absence of broader improvements in student performance likely reflects persistent instructional constraints and weak institutional incentives. Mediation

[4]For testing no-manipulation, I conduct the Cattaneo et al. (2020a) density test.

[5]Girls' secondary net enrollment in Bangladesh has remained below that of boys for the past two decades (World Bank, 2024). Completion rates also lag: by 2019, 31.6% of boys had completed secondary education compared to 27.3% of girls (World Bank, 2024). In contrast, completion is near universal in developed countries, where girls often outperform boys. See appendix figure A1.1a and A1.1b.

[6]Discussions related to multi-score regression discontinuity and related results are given in appendix B.

analysis shows no significant changes in teacher numbers or staff composition, suggesting that class sizes remained largely unchanged and limiting potential gains in instructional quality. Although the payroll subsidy reduced schools' direct salary obligations, the lack of data makes it difficult to determine whether the resulting fiscal space was redirected toward investments that could improve teaching quality or learning conditions. At the same time, low teacher quality, weak monitoring, low performance thresholds, and the absence of credible accountability mechanisms likely reduced incentives for both teachers and administrators to exert additional effort. These institutional weaknesses are compounded by broader governance challenges, including but not limited to weak functionality of school management committees, and persistent earnings gap between government teachers and MPO-supported teachers. To better understand the channels underlying the main results, the paper draws on descriptive evidence and institutional context, due to lack of school level data, to offer a plausible explanation for the absence of broad academic gains.

This study provides the first empirical evaluation of one of the largest and longest-running educational reforms in Bangladesh. The results connect the paper to a large body of literature highlighting the role of institutions, accountability, and governance in mediating the impact of education finance. Wößmann (2003) stresses institutional features as determinants of performance; while Kingdon and Teal (2010) and Suryadarma et al. (2006) link unionization and absenteeism in developing countries to weak learning highlighting how weak oversight, limited accountability, and bureaucratic inertia lead to governance failure. In this regard, Duflo et al. (2015) also shows that spending without accountability fosters inefficiency. The MPO subsidy in Bangladesh operates in the context of weak monitoring, political influence, and institutional corruption (World Bank, 2013), which altogether might have muted the effectiveness of the program. The absence of periodic evaluations of MPO schools, combined with a low threshold for continued support, may further reduce accountability and weaken incentives to improve school performance. Despite the payroll subsidy, I also do not observe any changes in staffing leaving class sizes plausibly unchanged. The absence of adjustments in teacher numbers and the lack of clarity regarding the reallocation of payroll savings may help explain the program's limited impact on student achievement. These challenges are consistent with the regional experiences: for instance, Chin (2005) documents how political rent-seeking in India undermined a teacher allocation reform, neutralizing its intended impact on class size or teaching quality.

This paper is also related to a broader literature on teacher compensation and student performance, which reports mixed evidence. Menezes-Filho and Pazello (2007) document a positive association between teacher wage increases and student proficiency in Brazil, whereas Muralidharan and Sundararaman (2011) show that higher paid civil service teachers in India often underperform relative to contract teachers. A common explanation is that sometimes the structure of compensation matters more than salary levels alone. Evidence

suggests that linking pay to performance can improve outcomes (Glewwe et al., 2010; Muralidharan and Sundararaman, 2011). In contrast, the MPO reform relied on unconditional salary payments that were not tied to student outcomes and offered limited mechanisms for teacher accountability. At the same time, it preserved a sizeable earnings gap between MPO teachers and government teachers in equivalent posts, potentially weakening motivation through perceived inequity and lack of accountability. Although performance based financing was proposed before in Bangladeshi MPO schools, it was abandoned because of political resistance (Asian Development Bank (2015)). Consistent with this, this study adds support to the claim that pay unrelated to performance do not improve educational attainment.

Finally, work on distributional impacts shows public educational spending raises achievement in disadvantaged districts in developed countries (Lafortune et al., 2018; Aaronson and Mazumder, 2011; Cascio et al., 2013), but effects in low-income regions, in a developing country context, are muted due to teacher labor-market frictions and governance constraints (Alderman et al., 2003; Marshall, 2009; Chin, 2005). In contrast to prior studies, however, I find no significant heterogeneity in aggregate performance gains across regions. It is important to note that the distinction between wealthier and less wealthy regions in this paper is based exclusively on variation across rural schools. Thus, "wealthier" regions should be interpreted as relatively better-off rural areas, rather than as wealthy regions in an absolute sense. Similarly, regions classified as more or less educationally corrupt capture relative differences in corruption exposure within the rural school sample. Taken together, this study adds an evidence on the effects of public financing on educational performances while examining potential differences along a regional poverty and educational corruption gradient.

The remainder of the paper is organized as follows. The next section provides the institutional background to the reform, followed by a discussion of the data sources, sample selection, and variables used in the analysis. The paper then sets out the identification strategy and empirical framework, covariate balance, and tests for non-manipulation, and subsequently reports the main estimation results, heterogeneity, and underlying mechanisms. The final section concludes with a summary of the findings and their implications.

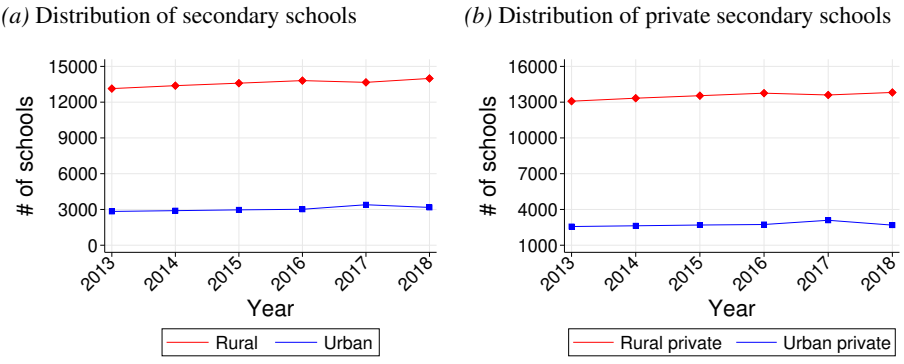
Institutional background

Some basic facts about secondary education in Bangladesh

Around 97% of secondary schools in Bangladesh operate in the private sector, the majority of which are located in rural areas. These rural schools are generally small in size and underperform relative to their urban counterparts, making quality improvement in this sector critical for reducing educational disparities across socio-economic groups. Between 2013 and 2018, over 13,000

secondary schools were located in rural areas, compared to about 3,000 in urban areas. This rural dominance, representing more than three-fourths of all schools, highlights the rural-centric nature of Bangladesh's secondary education system. Figure 1a presents the rural-urban distribution of schools during the pre-treatment period, while figure 1b shows the spatial distribution of private institutions. Together, the figures confirm that rural secondary education is overwhelmingly delivered by private providers, underscoring their pivotal role in expanding access.

Figure 1. Distribution of secondary schools



Notes: Panel (a) shows the total distribution of secondary schools across rural and urban areas between 2013 and 2018, while Panel (b) illustrates the distribution of private secondary schools over the same period. Data are disaggregated by rural and urban categories to highlight differences in school distribution patterns. Source: Bangladesh Bureau of Educational Information and Statistics (2013–2018).

In the context of Bangladesh, the urban–rural comparison of secondary schools can broadly be understood along two dimensions: institutional capacity and academic performance. Recent statistics indicate that urban schools maintained average enrollments of around 500 students, rising to nearly 700 by 2018, while rural schools' average enrollment stayed at a relatively lower level, from approximately 330 students in 2013 to 500 in 2018 (BANBEIS (2018)). Although this upward trend in both categories of schools likely reflects incremental improvements in infrastructure and the expansion of educational programs, the comparatively limited capacities of rural schools raise concerns regarding their long-term sustainability and ability to achieve economies of scale. While student–teacher ratios appear broadly similar across school types, performance outcomes (e.g., passing rate and share of distinctions in the public exam) reveal a far more pronounced disparity. Between 2013 and 2018, urban schools consistently outperformed rural institutions in secondary public examinations, with the performance gap widening over time (BANBEIS, 2018) likely due to spatial disadvantages, low teacher quality, limited incentives, and precarious employment conditions. This persistent gap in excellence compounds

broader disparities in overall achievement, thereby reinforcing socio-economic stratification within the education system.

Overall, several key patterns of secondary education in Bangladesh emerge: (i) the overwhelming presence of rural private schools highlights their centrality in Bangladesh's education architecture; (ii) rural schools consistently under perform across all metrics (e.g., enrollment, passing rates, and distinctions), making them the most vulnerable segment; (iii) the performance gap between urban and rural institutions is persistent and widening; (iv) given that most rural secondary schools are privately run and often outside the direct purview of government funding, direct government intervention may be essential. Without structural improvements in rural education, these disparities are likely to widen. The Bangladeshi government's emphasis on enhancing educational quality in privatized rural institutions is well-founded, as the data shows that these schools form the majority yet suffer the most.

The Monthly Pay Order (MPO) intervention

The Monthly Pay Order (MPO) is a government program in Bangladesh that provides salary support to teachers employed in approved non government secondary schools and colleges. Given that approximately 97 percent of secondary institutions are privately managed, the MPO constitutes a central instrument through which the state finances secondary education outside the public system. The MPO program is administered by the Directorate of Secondary and Higher Education (DSHE) under the Ministry of Education, Bangladesh. Through this scheme, the government directly disburses the basic salary of approved teachers according to the national pay scale applicable to public employees. Allowances and other benefits are only partially covered, so MPO teachers typically receive lower overall compensation than government school teachers but higher than non-MPO teachers in the rural areas. While MPO financing has made the scheme central to secondary education financing by reducing payroll uncertainty, the continued disparity in total compensation compared with government school teachers has fueled recurring dissatisfaction and collective action, with potential implications for teacher motivation and retention.

The program was introduced in 1981, initially covering 50 percent of basic salaries, and was gradually expanded to full coverage for teachers who satisfy regulatory requirements including certification by the Non-Government Teachers' Registration and Certification Authority (Asian Development Bank, 2015)[7].

[7]The establishment of the Non-Government Teachers' Registration and Certification Authority (NTRCA) is considered a key achievement of the Teacher Quality Improvement in Secondary Education Project (TQISEP), an Asian Development Bank aided government project, reportedly initiated under pressure to improve the quality of teacher recruitment. The NTRCA certification process has been described as successful, with significant reductions in rent-seeking behavior related to the registration of non-government schools and the certification of teachers. Notably, during the first three years of its operation, no cases of malpractice in teacher certification were reported.

Although concerns remain regarding the scope and pedagogical depth of the certification examination, the linkage between MPO eligibility and certification has strengthened standardization in recruitment and reduced informal influence in hiring practices.

The 2018 MPO reform and selection criteria

In 2018 the government initiated a major expansion of the MPO program through a competitive selection process aimed at incorporating additional non-government secondary schools into the payroll system. Prior to this reform, approximately 82 percent of non-government secondary schools were already receiving MPO support (Bhatta et al., 2019). The program is administered by the DSHE under the Ministry of Education. A public call for applications was issued in 2018, and schools submitted documentation through an online portal. Following administrative review and field verification by the Directorate of Secondary and Higher Education, the list of selected institutions was published in 2019. Performance information from the academic years 2015 to 2017 was used to construct eligibility scores. Salary disbursement to newly included schools began around 2020. For empirical purposes, 2021 is treated as the first full post reform year, as it represents the first complete academic year following implementation, despite disruptions associated with the pandemic. Moreover, no public examinations were conducted due to school closures in 2020, and students were promoted based on prior academic performance, making that year unsuitable for evaluating reform effects.

Selection criteria:

MPO eligibility was determined using a point based system that combined structural (basic) and performance related (crucial) criteria, as summarized in table 1.

Table 1. *Selection Criteria for Receiving MPO Subsidy*

Selection Criteria	Rules	Maximum Score
Panel A: Basic criteria		
(1) School tenure	5 points for every 2 years 25 points for 10 or more years	25
(2) # of enrolled students	300 for urban, 200 for rural (coeducation) 200 for urban, 100 for rural (girls only) 15 points for meeting minimum requirements 5 points increment for each 10% increase thereafter	25
Panel B: Crucial criteria		
(3) # of students appeared in public exam	60 for urban, 40 for rural (3 years average) 15 points for meeting minimum requirements 5 points increment for each 10 students thereafter	25
(4) Passing rate in public exam	70% (3 years average) 15 points for meeting minimum requirements 5 points increment for each 10% increase thereafter	25

Notes: Schools must attain the minimum score in each criterion to qualify for MPO inclusion. Source: Ministry of Education, Bangladesh (2018).

Basic criteria included school tenure[8] and minimum enrollment thresholds. Performance measures were based on three year averages of the number of students sitting for the grade 10 public examination and the corresponding passing rate over 2015 to 2017. Schools were required to meet minimum thresholds in each category to qualify for inclusion (Ministry of Education, Bangladesh (2018)). Because public examinations are centrally administered and externally graded, and because eligibility scores rely on predetermined multi year averages, institutions had limited scope to manipulate reported outcomes.

Once granted, MPO status is not subject to periodic reapplication, provided that schools continue to comply with regulatory requirements and maintain approved staffing positions. Schools that were not selected continue to operate as privately managed institutions financed primarily through student fees and local contributions. In the absence of MPO support, teacher salaries are generally lower and less stable, as they depend on school level revenue. In practice, both MPO and non-MPO schools increasingly prioritize recruitment of teachers certified by the Non Government Teachers Registration and Certification Authority. For non-MPO institutions, appointing certified teachers strengthens their prospects of future inclusion in the MPO scheme, as certification is a key eligibility requirement. At the same time, teachers have strong incentives to obtain NTRCA certification, since it enhances their eligibility for positions in MPO supported schools and improves their prospects for more stable, government financed employment. While MPO funding reduces school's fiscal pressure and may create scope to moderate tuition fees, the extent of such adjustments varies across institutions. The counterfactual in this study consists of comparable non-government schools that applied but did not receive MPO inclusion during the reform wave. These schools operate under similar regulatory conditions but without government financed salary support.

The basic criteria represent the dimensions in which the majority of schools readily surpass the threshold score. Consequently, I treat these two criteria as fixed in the analysis; meaning all schools in the sample have passed the thresholds related to basic criteria. The identification strategy therefore exploits variation in the two crucial performance-related criteria, which are discussed in greater detail later.

Is MPO important for schools?

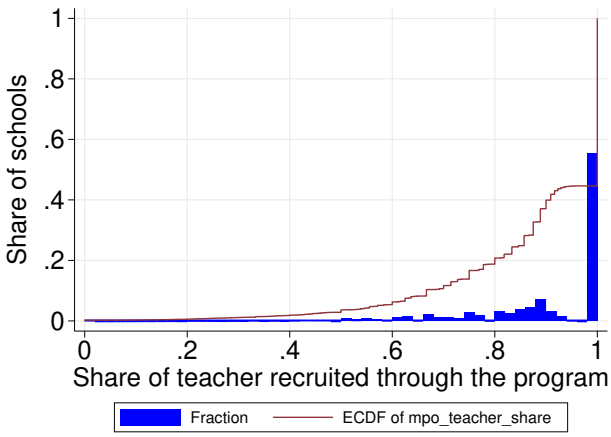
The MPO program is extensive in scope. At the secondary and higher secondary levels, a large share of non-government institutions in Bangladesh are MPO-enlisted. According to the Directorate of Secondary and Higher Education,

[8]Here, "school tenure" does not refer to the tenure from establishment date. In Bangladesh, non-government secondary schools must first register with the Ministry of Education to be considered for MPO financing. The Ministry reviews these applications against a set of administrative criteria, and the tenure count begins once the registration is formally approved.

by 2017 the MPO scheme covered a substantial portion of non-government secondary schools receiving government salary support (Bhatta et al., 2019). This widespread reach underscores the program's central role in sustaining the private education sector, which complements state provision in meeting national demand for secondary education.

Given that teacher salaries account for more than 90% of education budgets in both government and non-government secondary schools, payroll is by far the largest component of financing. For non-government institutions, which make up a substantial share of the secondary system, these costs are predominantly met through the MPO program, financed from the government's revenue budget. Teachers in MPO schools are paid directly from the treasury via electronic transfers, reducing reliance on private tuition income and buffering schools against enrollment fluctuations. In this way, the program acts as a stabilizing force in a sector often characterized by volatile funding and limited resilience. Importantly, the MPO scheme is designed as a partnership rather than direct government control: while the state covers payroll expenses through long-term, recurring subsidies, day-to-day management, recruitment, and pedagogy remain the responsibility of school administrations. This arrangement allows local flexibility while ensuring that public resources sustain essential educational services and institutional viability. Beyond providing salaries, MPO status is widely regarded as a marker of credibility and legitimacy. For many non-government schools and colleges, enlistment enhances community standing, facilitates teacher recruitment, and secures financial stability that supports broader institutional operations.

Figure 2. Distribution of payroll coverage



Notes: The figure illustrates the staffing composition of teachers in MPO recipient private secondary schools. This histogram is based on data from 2016. The horizontal axis reflects the relative share of MPO-subsidized employees within educational institutions, and the vertical axis shows the fraction of schools relative to the MPO share.

Recent data show that MPO subsidies cover the full payroll in more than half of recipient schools, while over 90% receive support covering at least two-thirds of payroll expenses. Figure 2 illustrates this heavy reliance: the histogram shows a large mass of schools recruiting nearly all teachers under MPO, with the cumulative distribution function confirming that most institutions achieve very high levels of MPO-supported staffing. These patterns highlight the program's centrality to the operational sustainability of Bangladesh's secondary schools, particularly in maintaining teaching staff and ensuring continuity in service delivery. Without MPO funding, the ability of non-government schools to function effectively would be severely constrained, jeopardizing access to education for millions of students.

Challenges of MPO intervention

Over the years, MPO has become deeply embedded in Bangladesh's broader education financing strategy and continues to play a vital role in promoting access, continuity, and basic financial equity for teachers outside the formal public service. However, in practice, the linkage between MPO subsidies and performance outcomes has been less robust than originally envisioned. While official policy criteria exist for maintaining MPO status, the system has faced challenges in consistently enforcing them. Schools are rarely removed from the MPO list for under-performance, and political considerations often influence their continuity. Although school management committees (SMCs) were intended to strengthen local governance, their effectiveness has varied, raising concerns about the quality of oversight. Moreover, the program's low performance cutoffs may have weakened incentives for effort, since teachers and administrators could secure funding by merely meeting minimal requirements. A World Bank report (2013) finds that subsidies as a policy tool to enhance educational quality and accountability have largely failed, owing to weak incentives for maintaining performance, as underperforming schools face few penalties. Political support, rather than adherence to quality standards, frequently determines the continuation of MPO status, leaving many affiliated schools below expected performance levels. While proposals for performance-based financing were considered, they were never implemented, partly due to the challenges of stakeholder consensus (Asian Development Bank, 2015). Nonetheless, opportunities remain to align MPO disbursements more closely with educational outcomes, including thorough transparent reporting, stronger accountability mechanisms, and greater public awareness of school performance.

Data and sample

The primary dataset for this study is drawn from the Bangladesh Bureau of Educational Information and Statistics (BANBEIS). This confidential dataset provides school-level data on institutional characteristics, student performance,

and teacher profiles from 2012 to 2023 for the universe of secondary schools in Bangladesh. It also includes detailed information on the 2018 MPO application round, covering scores on each selection criterion as well as the final nomination outcomes for all applicant schools. In that year, 1,839[9] rural secondary institutions (grades 9–10) applied for MPO funding, of which 671 schools were nominated for MPO financing.

To strengthen internal validity and ensure comparability, I restrict the analysis to schools that satisfy the core eligibility conditions of the MPO selection process. First, I retain only schools meeting the tenure requirement. As noted earlier, tenure refers not to the time since establishment, but to the length of time a school has been formally recognized through the administrative process relevant for MPO inclusion (see footnote 7). This restriction ensures that schools in the estimation sample have sufficient pre-reform administrative history to be meaningfully assessed under the scoring system. Second, I restrict the sample to schools that satisfy the "number of enrolled students" (second basic criterion). Since more than 90 percent of schools meet this condition,[10] using it as a running variable in a separate regression discontinuity design would leave too few observations below the cutoff, particularly near the threshold, thereby reducing the precision of the estimates. The resulting sample, after fixing these two criteria, comprises 1,509 schools, including 671 treated and 838 untreated schools.

The school level data are further augmented with subdistrict (local regional) level poverty indicators to capture the influence of local socioeconomic conditions on educational outcomes. These indicators are measured using poverty headcount ratios derived from the most recent nationally representative household survey conducted by the Bangladesh Bureau of Statistics in 2022. Defined as the share of individuals living below the national poverty line, these ratios serve as a proxy for local economic deprivation. In addition, the analysis incorporates education sector corruption indicators from the 2021 National Household Survey conducted by Transparency International Bangladesh, as well as MPO related protest event data from the GDELT 2.1 Global Event Database[11], to contextualize institutional quality and political dynamics. Merging these measures with the school dataset enables an examination of whether

[9]The analysis is restricted to co-educational secondary schools, which is the most common school type in rural areas. This restriction is warranted on both institutional and econometric grounds. Girls-only schools were assessed under a distinct set of selection criteria, rendering them not directly comparable within the same regression discontinuity design. Furthermore, the number of girls-only schools receiving MPO support is very small, leaving too few treated observations to permit sufficiently precise statistical inference.

[10]In total, 1,663 of the 1,839 schools satisfied the enrollment criterion for MPO selection.

[11]GDELT (Global Database of Events, Language, and Tone) is a large-scale, publicly available dataset that automatically codes political, social, and economic events from global news coverage using natural language processing and the CAMEO event taxonomy (Leetaru and Schrodt, 2013). The database provides structured, event-level information derived from media reports and is updated at high frequency across countries and languages.

the effectiveness of MPO financing varies systematically across differences in poverty, governance conditions, and local accountability environments, thereby providing a broader assessment of the program's distributional and institutional implications.

The empirical analysis examines four categories of outcome variables constructed from BANBEIS administrative records.: overall educational performance, secondary school participation rate, girls' performance, and staffing. Each category is precisely defined in the subsequent sections.

- *Passing rate:* The proportion of students enrolled in a school who successfully passed the nationally administered secondary examinations in a given year.
- *Distinction rate:* The share of students achieving the highest grade (GPA 5.0) in the secondary examinations, expressed as a fraction of all examinees. This serves as a proxy for academic excellence at the tail of the achievement distribution.
- *Secondary school participation (overall):* Secondary school participation of all students. This reflects the aggregate participation at in secondary education.
- *Girls' secondary school participation ratio:* The ratio of girls-to-boys secondary school participation rate. This indicator captures relative gender participation in schooling and is used to assess equity outcomes.
- *Girls' passing rate:* The proportion of girls enrolled who passed the secondary examinations, measured separately from the overall passing rate to assess gender-specific learning outcomes.
- *Girls' distinction rate:* The share of girls attaining the highest grade (GPA 5.0). This isolates high academic achievement among girls, complementing the broader distinction rate.
- *Total teachers:* The total number of full-time teachers employed in each school. This indicator reflects the staffing capacity available for instruction.
- *Share of female teachers:* The number of female teachers divided by the number of total teachers in a school. This variable captures the gender composition of teaching staff and may influence both gender-sensitive pedagogy and the perceived safety of schooling environments for girls.

Econometric framework and identification

Identification

Selection into the MPO scheme is governed by a rule-based scoring mechanism that evaluates schools along four dimensions: school tenure, number of enrolled students, average number of students appearing in public examinations (averaged over three years prior to the reform), and the corresponding average

passing rate over the same period. Each criterion carries a designated cutoff score, and a school must meet or exceed the minimum threshold in all four dimensions to qualify for MPO inclusion. This creates a treatment assignment mechanism in which eligibility is determined by fixed thresholds, naturally lending itself to a regression discontinuity design (RDD). Under the assumption that schools cannot precisely manipulate their scores to selectively cross the cutoffs, the discontinuity in eligibility at the thresholds provides a credible basis for causal inference.

As explained in detail in the previous section, the sample is restricted by fixing two basic eligibility criteria that are of limited relevance for identification. This improves comparability and allows the empirical design to focus on the margins that are most informative for the causal estimation. The key criteria are the three-years average of the number of students appearing in public examinations and the corresponding passing rate, both measured using pre-treatment data. These criteria are binding for a substantial share of schools, and eligibility is sharply determined by whether a school meets or exceeds the thresholds of at least 40 appearing students and a 70 percent passing rate, conditional on satisfying the other criteria at a time. Variation around these cutoffs forms the basis of the identification strategy. Specifically, the analysis implements a sharp regression discontinuity design that compares schools located just above and just below the relevant eligibility thresholds and are therefore plausibly similar in all other respects. Because treatment assignment changes discretely at these performance cutoffs, the design approximates random assignment within a narrow neighborhood of the threshold, yielding locally unbiased estimates of the causal effect of MPO funding. The presence of these two key running variables also makes it possible to implement the multi-score regression discontinuity design proposed by Cattaneo et al. (2020b) in two-dimensional treatment-control space, which results in a continuum of treatment effects defined along the boundary of the treatment region, provided there is sufficient support (see appendix B).

The final sample in this study consists of 1,509 schools, with 671 in the treated group and 838 in the control group. The sample is well balanced across treatment status, ensuring comparability between schools just above and just below the eligibility thresholds. Although the forcing variables are discrete, their granularity is sufficient to support credible estimation within a regression discontinuity framework. This design facilitates a transparent and internally valid evaluation of the MPO scheme by exploiting the policy-induced discontinuity in eligibility to isolate its causal impact on school performance.

Empirical framework

The identification strategy, in this paper, exploits discontinuities in two running variables: (i) average passing rate, and (ii) the average number of students appearing in public examinations, and attempts to estimate the causal effect

of MPO financing on school outcomes around the pre-specified eligibility thresholds of these running variables. The empirical model is specified as follows:

$$Y_i = \alpha + \tau D_i + \sum_{k=1}^{p} \gamma_k (R_{ij} - c_j)^k + \sum_{k=1}^{p} \lambda_k D_i \cdot (R_{ij} - c_j)^k + \varepsilon_i, \quad j = 1,2 \quad (10)$$

where Y_i denote the outcome for school i, and D_i the treatment assignment, where $D_i = 1$ if $(R_{ij} - c_j) \geq 0$ for all j. Here, R_{ij} represents the raw score for school i on running variable $j \in \{1,2\}$, and c_j is the corresponding cutoff. The running variables are re-centered at their cutoffs, $(R_{ij} - c_j)$. I estimate two sharp RD designs separately, one for each running variable, thereby generating two distinct treatment effect estimates. The general specification allows for different polynomial orders of the control function, with slopes permitted to vary on either side of the cutoff by interacting the treatment indicator with the polynomial terms. In this paper, however, I rely on a parametric linear control function. As a robustness check, I also implement the non-parametric multi-score regression discontinuity approach of Cattaneo et al. (2020b), where the identifying variation is effectively driven by the key running variables: the three-year averages of the number of students appearing in public examinations and the corresponding passing rate. This yields a multi-score RD design in which treatment status changes along a boundary in the two-dimensional score space. See appendix B for detailed discussion and results. Due to limited information on school characteristics, covariates are excluded from the all specifications. Nonetheless, I conduct balancing tests on predetermined outcomes. Robust standard errors are used in all the results.[12]

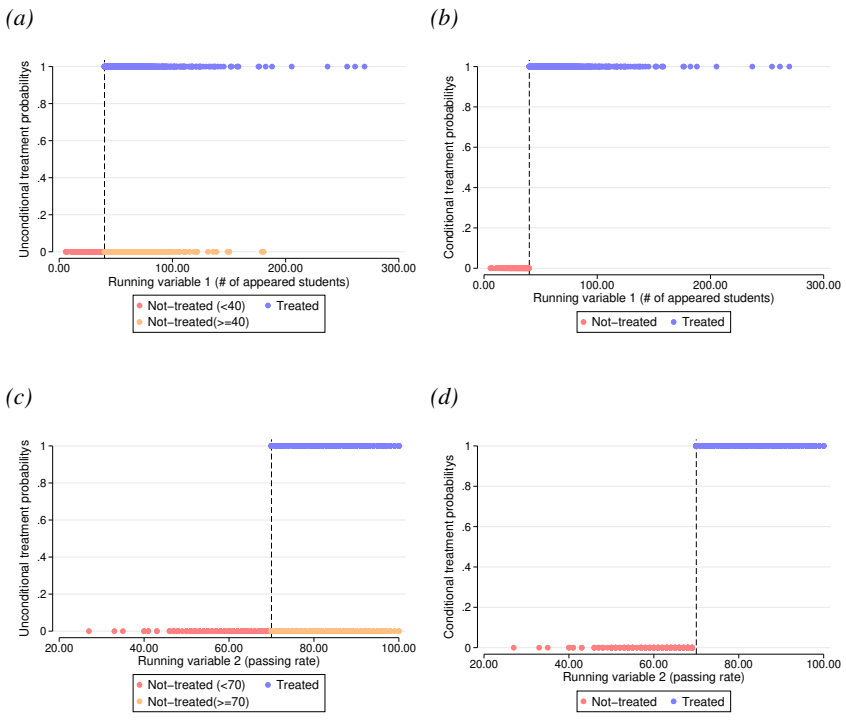
The major parameter of interest is τ that captures the treatment effect of receiving the MPO funding, under the assumption that there is no precise manipulation of the running variables at the cut-offs. This assumption ensures that units just above and below the threshold are comparable, allowing the cutoff to mimic random assignment. If schools can sort precisely around the threshold, for instance by bunching just above it to secure treatment, the continuity assumption is violated and the estimated treatment effect may be biased. To assess this concern, I use graphical diagnostics and McCrary (2008)-style density tests developed by Cattaneo et al. (2020a), which provide indirect evidence on manipulation by testing for discontinuities in the density of the running variable at the cutoff. In addition, I evaluate identification by re-estimating equation (1) using predetermined outcomes measured before the MPO expansion. The absence of discontinuities in these baseline variables supports the validity of the research design. Finally, it is important to distinguish between sharp and fuzzy RDD designs in this context as it has two running variables and assignment to treatment depends on both. The main empirical analysis in this

[12] See Kolesár and Rothe (2018) for a discussion of inference under discrete running variables.

paper relies on single-cutoff sharp RDDs, with each specification using one running variable while conditioning on crossing the threshold associated with the other. In this conditional case, treatment assignment is perfectly determined by whether the running variable, in question, crosses the cutoff: $D_i = \mathbb{1}(R_{ij} \geq c_j)$ for $j \in \{1,2\}$.

Treatment assignment, covariate balance and bunching

Figure 3. Treatment assignment mechanism based on running variables



Notes: The figures illustrate the treatment assignment probability (first stage) based on two running variables: (a) the average number of appeared students, and (b) the average passing rate. The left panels show unconditional treatment probabilities, where assignment appears fuzzy because some schools cross one threshold but remain untreated (orange dots) because of non-attainment of another threshold. The right panels show conditional treatment probabilities, where treatment is evaluated conditional on crossing both thresholds, producing a sharp discontinuity at the cutoff. Treated schools are shown in blue, while non-treated schools are in red (below the threshold) or orange (crossing one threshold only).

This section outlines the treatment assignment mechanism and examines the validity of the regression discontinuity (RD) design through two key diagnostics. First, I describe how schools are classified as treated or untreated based on whether they meet the eligibility thresholds on the crucial criteria. Second, I test whether predetermined covariates are balanced at the cutoff,

ensuring comparability between treated and untreated schools. Third, I assess the possibility of manipulation or bunching in the running variables around the eligibility threshold, which could compromise the design's internal validity. Taken together, these elements both clarify the assignment of treatment and establish the credibility of the RD framework before turning to estimates of program effects.

Treatment assignment based on running variables

The treatment assignment mechanism used in this paper is illustrated using two running variables: the passing rate and the number of appeared students. Figures 3a and 3c show the unconditional treatment assignment, where treatment probability is defined without conditioning on the other running variable. In this case, assignment is fuzzy, as some schools to the right of a threshold remain untreated (solid orange dots), reflecting that they cross only one of the two thresholds. In contrast, figures 3b and 3d depict the conditional treatment assignment, where treatment probability is evaluated conditional on crossing both thresholds. This conditional assignment sharpens the rule, producing a clear discontinuity in treatment status at the cutoff. As stated earlier, I adopt the sharp RD approach in this paper by using each running variable separately and conditioning on the other at the same time, so that units below the cutoff are never treated and those above are always treated. This deterministic assignment rule generates a credible quasi-experimental setting, as units just above and below each cutoff can be considered comparable in expectation. The validity of the design hinges on the continuity of potential outcomes around the threshold, ensuring that any observed discontinuity can be attributed to treatment rather than to underlying trends.

Covariate balance

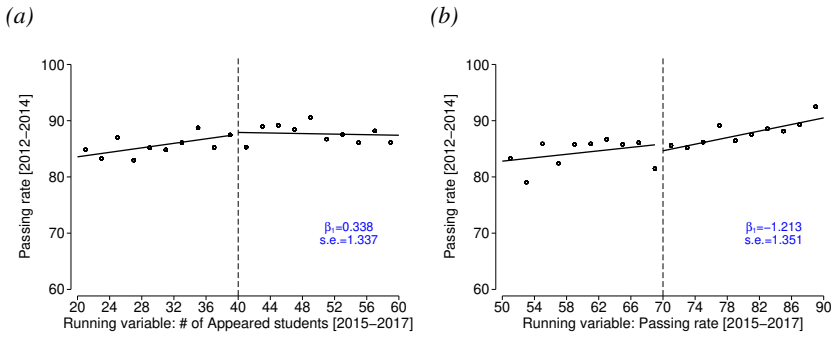
A core requirement for the credibility of a regression discontinuity design is the absence of systematic differences in predetermined characteristics around the eligibility cutoff. If schools just above and just below the threshold are similar in attributes that cannot be influenced by the policy, such as prior academic performance, gender composition, or staffing, then treatment assignment near the cutoff can be interpreted as quasi-random. Conversely, any discontinuity in these covariates would undermine comparability and weaken the design's internal validity. This subsection evaluates whether baseline characteristics vary smoothly across the cutoff, providing a key diagnostic for the RD framework.

Student performance:

Figures 4a and 4b plot RD estimates for pre-treatment average passing rates. The estimated discontinuities, 0.338 and -1.213 for the two running variables, are small and statistically insignificant, indicating no systematic differences between schools just above and just below the cutoff. Pre-treatment share

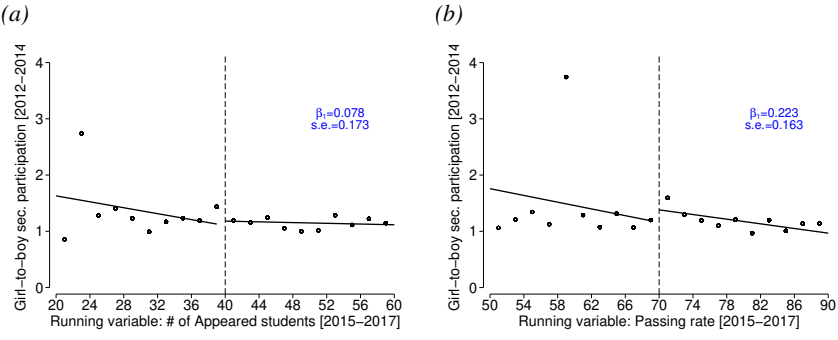
of distinctions further confirm this balance: coefficients of 0.001 and –0.003 (figures A2.1a and A2.1b in the appendix) show no meaningful gap in the share of students earning distinctions. These results suggest that high-performing students were not systematically sorted across treatment and control schools prior to the allowance, reinforcing the comparability of schools near the threshold.

Figure 4. Covariate balance: Pre-determined average passing rate



Notes: The figure plots the pre-treatment average passing rate against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. Pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (two-point bins), and the fitted regression lines are obtained from Equation (10). Estimated discontinuities and standard errors are reported. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure 5. Covariate balance: Pre-determined secondary school participation by girls



Notes: The figure plots the pre-treatment girls-to-boys secondary school participation ratio against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero. Pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (two-point bins), and the fitted regression lines are obtained from Equation (10). Estimated discontinuities and standard errors are reported. * $p < .1$, ** $p < .05$, *** $p < .01$.

Secondary school participation:

Gender balance at the secondary level shows no evidence of discontinuity at the threshold. The girls-to-boys secondary participation ratio yields treatment

coefficients of 0.078 and 0.223 across the two running variables (figures 5a and 5b), and both are statistically insignificant implying that treated and untreated schools were similar in gender composition prior to the allowance. These results suggest that observed differences are likely due to random variation rather than systematic imbalance.

Girls' performance:
Appendix figures A2.2a and A2.2b show RD estimates for girls' passing rates. The discontinuities (–0.156 and –1.581) are negative but statistically insignificant, providing no evidence of systematic pre-treatment differences. Appendix figures A2.3a and A2.3b confirm that girls' distinctions also exhibit no discontinuity. Together, these findings indicate that academic performance for girls was balanced across schools near the threshold, lending further support to the validity of the design.

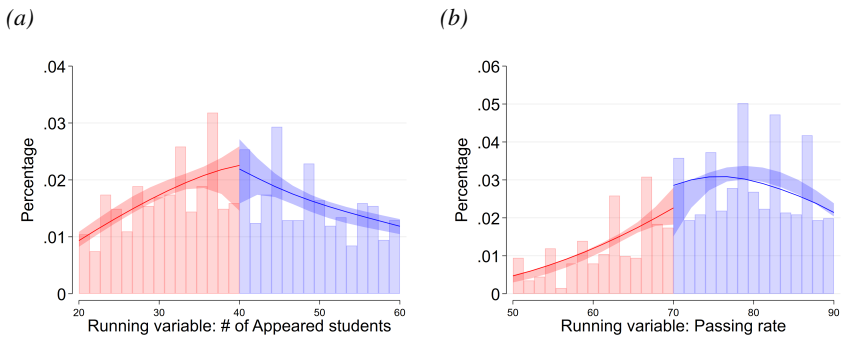
Staffing levels:
Number of teachers also display no evidence of imbalance. Estimated discontinuities of –0.340 and –0.139 across the two running variables (appendix figures A2.4a and A2.4b) are small and statistically insignificant. This smooth distribution of staffing across treated and untreated schools suggests that post-policy changes in teacher numbers can be credibly attributed to the allowance rather than pre-existing differences.

Bunching

To evaluate the potential for manipulation around the cutoff, I employ the Cattaneo et al. (2020a) density discontinuity test [13], which examines whether the frequency of observations changes significantly at the threshold. Two versions of the test were implemented, one for each running variable:

[13]The Cattaneo et al. (2020a) test is a formal statistical procedure designed to detect manipulation of the running variable in Regression Discontinuity Designs (RDD). Specifically, it assesses whether units are disproportionately concentrated just above or below the treatment cutoff—an issue known as "bunching"—which could compromise the validity of the RDD by violating the assumption of random assignment around the threshold. Unlike earlier density tests, this method uses local polynomial density estimation to compare the distribution of the running variable on either side of the cutoff in a fully data-driven way. The test statistic evaluates whether there is a discontinuity in the estimated density at the threshold. A statistically significant result would suggest that the running variable may have been strategically manipulated, undermining the as-good-as-random nature of the design near the cutoff.

Figure 6. Test for non-manipulation



Notes: The figure presents the manipulation test for bunching using the density discontinuity test proposed by Cattaneo et al. (2020a). Panel (a) uses the number of appeared students as the running variable, and Panel (b) uses the passing rate. The histograms show the distribution of the running variables, while the solid lines with shaded confidence bands display the estimated densities from local polynomial regressions. The procedure tests for discontinuities in the distribution of the running variable at the cutoff, which would be consistent with bunching or manipulation. Equal bandwidths on both sides of the cutoff have been chosen for the test; however, the results are also robust to using the data-driven bandwidth selectors proposed by Cattaneo et al. (2020a). Confidence bands are constructed following the robust inference procedures by Cattaneo et al. (2020a).

Table 2. *Bunching Test Results (Cattaneo, Jansson, and Ma, 2020)*

Running Variable	Test Statistic	p-value
Running Variable 1 (# of appeared students)	0.050	0.960
Running Variable 2 (passing rate)	-0.940	0.347

Notes: The table reports the test statistics from the manipulation test for bunching based on the density discontinuity test of Cattaneo et al. (2020a). The running variables are (1) the number of appeared students and (2) the passing rate. This specification uses a second-degree local polynomial order to construct the density estimators. The test statistics and corresponding p-values are shown. Higher p-values indicate that the null hypothesis of no discontinuity in the density at the cutoff cannot be rejected, suggesting no evidence of manipulation or bunching around the threshold. Equal bandwidths on both sides of the cutoff were used for the test, and the results are robust to the data-driven bandwidth selectors proposed by Cattaneo et al. (2020a).

In both cases, the null hypothesis of no discontinuity in the density of the running variable at the cutoff cannot be rejected (figure 6a, 6b and table 2). These results imply that schools did not manipulate their assignment status to qualify for the allowance. The findings are robust to alternative functional form specifications, as shown in the appendix table (table A3.1). This is a particularly informative result given the context of weak monitoring. The lack of bunching might indicate institutional inertia or limited capacity among schools to exploit the low eligibility thresholds. It also enhances the credibility of the RDD design, ensuring that treatment assignment can be considered as-good-as-random around the cutoff.

Does MPO affect school outcomes?

The section examines whether the program program generated measurable changes in student achievement, participation, and other school-level outcomes.

Table 3. *Regression Discontinuity Estimates*

	Running Variable 1	Running Variable 2
Panel A: Student Performance		
Passing rate	1.126	0.349
	(1.124)	(1.268)
Rate of distinction	0.010	0.002
	(0.007)	(0.005)
Panel B: Secondary school participation		
Overall	0.001	-0.003
	(0.005)	(0.008)
Girls	0.212**	0.151*
	(0.089)	(0.083)
Panel C: Performance for girls		
Passing rate	0.876	1.838
	(1.723)	(2.459)
Rate of distinction	0.013	-0.003
	(0.009)	(0.006)
Observations	709	757

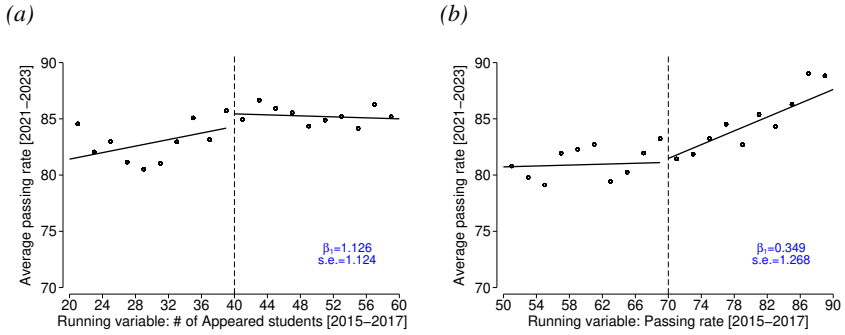
Notes: Each cell reports the RD coefficient estimate at the cutoff based on Equation (10). Robust standard errors are shown in parentheses. Both running variables (number of appeared students and passing rate) are re-centered so that the cutoff equals zero. Significance levels are denoted as follows: * $p < .1$, ** $p < .05$, *** $p < .01$.

Student performance:

Table 3, figure 7a and 7b present the estimates from the RD regression, with coefficients of 1.126 and 0.349 for the two running variables, respectively. In post-treatment period, no discontinuity is observed at the cutoffs and the coefficients are statistically insignificant. This pattern suggests that the payroll program, on average, may not have been effective in accelerating the academic performance of the treated schools than before. The small and statistically insignificant treatment effects imply that schools did not translate the additional funding into measurable gains in instructional quality or student performance. A plausible explanation is that, given the minimal accountability framework, the subsidy was perceived more as a guaranteed entitlement than as a reward linked to improved academic outcomes. This result goes in the spirit of of the findings of Muralidharan and Sundararaman (2011), who document that spending unconditional money on additional school input do not bring about academic performance of the students. Importantly, these findings suggest that injecting financial resources without complementary measures for accountabil-

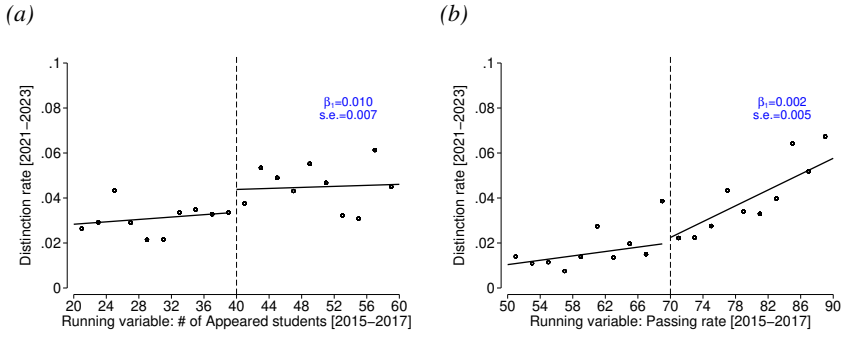
ity or targeted support may yield limited pedagogical returns.[14] It is, however, important to note that there might be other forces that drive this result.

Figure 7. Outcome: average passing rate



Notes: The figure plots the post-treatment average passing rate against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure 8. Outcome: average distinction rate

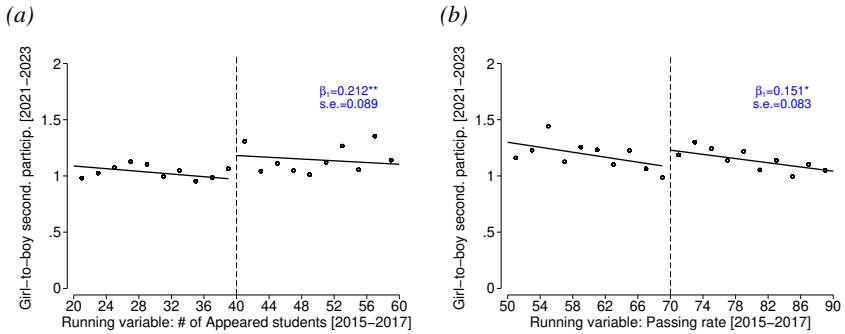


Notes: The figure plots the post-treatment average distinction rate against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

[14] In some settings, performance-based incentives may have limited or even no impact. Goodman and Turner (2013) evaluate a group-based teacher incentive program in New York City and report that it had little impact on student achievement.

This allowance program also appears to have had no measurable effect on distinction rates,[15] on average for the schools (figure 8). Coefficients from the RD estimates, presented in figure 8, are close to zero—0.010 and 0.002 for the two running variables—and statistically insignificant. This lack of change suggests that the policy fails to foster a significant culture of academic excellence or to motivate schools to move students into higher achievement tiers. If distinctions serve as indicators of instructional quality or rigorous academic preparation, their stagnation post-policy highlights the program's limited influence on pedagogical shift, possibly due to the absence of administrative pressure or incentives.

Figure 9. Outcome: Secondary school participation for girls



Notes: The figure plots the post-treatment average girls-to-boys secondary school participation ratio against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

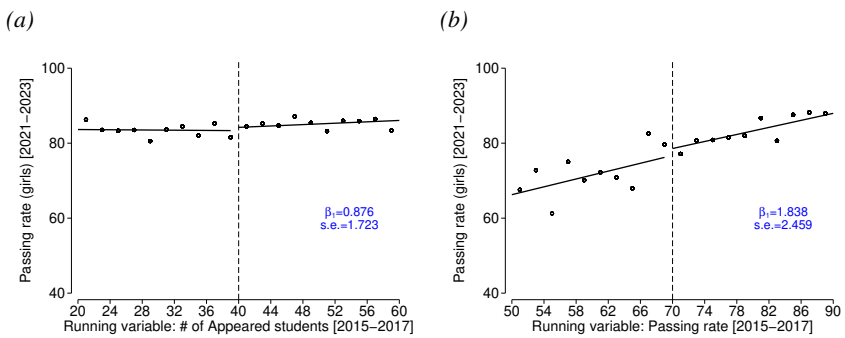
Secondary school participation and girl's performance:
The secondary school participation for girls at the higher secondary level shows a positive and weakly significant change following the policy. The estimated coefficients are 0.212 and 0.151 for the two running variables, significant at the 5% and 10% levels, respectively, as shown in figure 9a and 9b. These values indicate that, for treated schools, now approximately 0.21 and 0.15 more girls participate in the secondary school education, along the two running variables respectively. Results from the multidimensional RDD further corroborate this pattern (appendix table B1), pointing to modest gains in girls' participation within the two-dimensional treatment–control space, particularly among smaller, high-performing schools located close to both eligibility thresholds. Together, these findings suggest that the payroll subsidy program may

[15]The distinction rate is the share of students receiving the highest honors, among all who appeared at the end-of-the school public exam.

have little impact on gender balance at the targeted level of education. One plausible mechanism is that MPO schools may charge lower effective fees due to government salary support, thereby reducing the financial burden on households with school-aged girls. Even where formal fees remain similar, the perception that the state bears part of the cost may increase parental willingness to enroll daughters. Moreover, MPO status confers institutional legitimacy and subjects schools to central oversight, potentially signaling a safer and more stable academic environment for girls. However, the evidence is relatively weak, and the result should therefore be interpreted with caution.

I also try to disentangle the overall school performance by zooming in on to girl students. Figure 10a and 10b report the treatment coefficients for girls' passing rates, estimated at 0.876 and 1.838 for the two running variables, respectively, though neither is statistically significant. Additionally, I also do not see any significant discontinuity for girls' distinction rate (figure 11). These results point that treated schools do not experience meaningful improvements in girls' academic performance either. From a policy perspective, the program fails to improve the most basic student outcome.

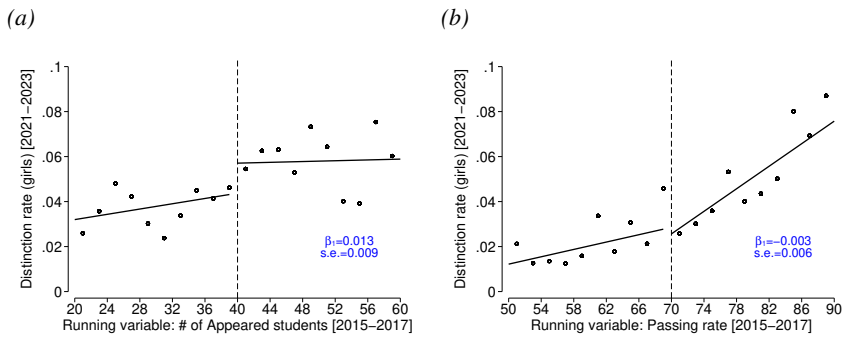
Figure 10. Outcome: average passing rate for girls



Notes: The figure plots the post-treatment average passing rate for girls against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Overall, the regression discontinuity estimates indicate that the teacher payroll subsidy program produced limited measurable impacts across most educational outcomes. For test scores and distinction rates, the estimated effects are small and statistically insignificant, and in some cases close to zero, suggesting no change in academic achievement. These overwhelmingly null findings raise important questions about the design and implementation of the scheme: while financial incentives can improve educational delivery, their effectiveness depends on conditionality, monitoring, and accountability. In the

Figure 11. Outcome: average distinction rate for girls



Notes: The figure plots the post-treatment average distinction rate for girls against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

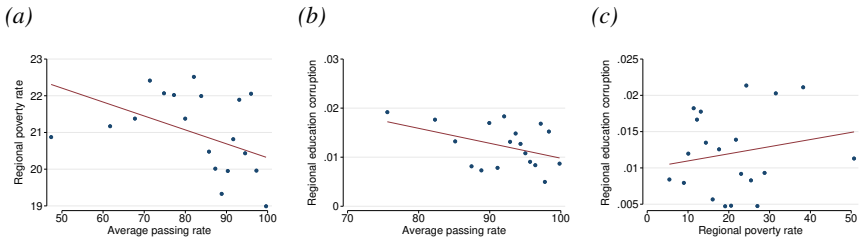
case of Bangladesh, the program's lenient thresholds and lack of oversight may have sent a weak signal to schools regarding the intended behavioral change, resulting in the subsidy being absorbed as a fixed income supplement rather than leveraged to improve quality or efficiency. The notable exception is girls' participation at the higher secondary level, where the girls-to-boys enrollment ratio at the secondary level rose by 0.212 and 0.151 along the two running variables, significant at the 5% and 10% levels, respectively. However, this result should be considered with caution. Building on these results, it is important to consider whether regional socio-economic conditions or educational corruption influenced the program's impact. Educational infrastructure, baseline gender gaps, and administrative quality can vary substantially across regions, shaping both how the subsidy was utilized and the extent to which it translated into improved outcomes.

Do regional socioeconomic conditions and corruption matter for school performance?

Regional variation in socioeconomic conditions and governance quality may shape the effectiveness of educational spending. A broad literature shows that the returns to education subsidies are often heterogeneous across local contexts, and that public spending may be less effective in poorer areas where institutional and administrative constraints are more severe (Alderman et al., 2003; Marshall, 2009). More generally, empirical work has long documented a close association between income levels and governance quality, with weaker institutions tending to be concentrated in lower-income settings (Acemoglu et al., 2001; Kaufmann

et al., 1999). In the context studied here, these considerations are particularly relevant given the focus on rural secondary schools, where both economic disadvantage and governance weaknesses may plausibly limit the ability of additional public resources to translate into improved learning outcomes.

Figure 12. Regional poverty, educational corruption, and school outcome



Notes: Figure 12a presents a binscatter plot of the relationship between average regional poverty headcount ratio and average school passing rates. The vertical axis reports the regional poverty headcount score, while the horizontal axis reports the mean passing rate at the school level. Each point represents the average outcome within evenly spaced bins of the mean passing rate, and the solid line denotes the fitted linear regression. Figure 12b shows the binscatter relationship between regional poverty headcount rates and average corruption in education. Solid lines indicate linear fits estimated using ordinary least squares. Poverty data are obtained from the Poverty Map of Bangladesh 2022 (BBS and World Bank), while school performance and corruption measures are constructed from administrative education records. Finally, Figure 12c shows the binscatter relationship between average regional corruption in the education sector and average school passing rates. The vertical axis reports the regional corruption index, while the horizontal axis reports the mean passing rate at the school level. Each point represents the average outcome within evenly spaced bins of the mean passing rate, and the solid line denotes the fitted linear regression.

Motivated by these considerations, the following analysis examines whether the effects of MPO funding vary systematically with regional poverty and corruption in the education sector. To do so, the school-level data are linked to spatially disaggregated poverty headcount rates from the Poverty Map of Bangladesh 2022, jointly produced by the Bangladesh Bureau of Statistics and the World Bank, as well as regional indicators of education-sector corruption.[16] The descriptive evidence reveals substantial geographic variation in both dimensions. Moreover, the binscatter relationships in figure 12 indicate that regional poverty is negatively associated with school passing rates and positively associated with corruption in education, suggesting that poorer regions tend to exhibit both weaker educational performance and lower governance quality. These

[16]Education-sector corruption is measured as the average number of education-related incidents of bribery, non-cooperation, or denial of services reported per household within a given local region. The measure is constructed by aggregating household-level reports of these corruption-related experiences and expressing them as a regional average, so that higher values indicate a greater prevalence of corruption and service irregularities in the local education sector.

descriptive patterns are consistent with the broader view that socioeconomic disadvantage and institutional weakness are closely intertwined.

Table 4. *Regression Discontinuity estimates by poverty- and corruption-level*

	Poverty level				Education corruption level			
	Poor region		Rich region		More-corrupt region		Less-corrupt region	
	RV1	RV2	RV1	RV2	RV1	RV2	RV1	RV2
Panel A: Student Performance								
Passing rate	0.622	1.977	1.974	-0.946	3.822	-0.124	0.094	-1.400
	(1.547)	(1.830)	(1.520)	(1.647)	(2.591)	(3.784)	(1.346)	(1.374)
Rate of distinction	0.001	0.003	0.023*	0.002	0.018	-0.022	0.007	0.001
	(0.008)	(0.006)	(0.014)	(0.009)	(0.022)	(0.016)	(0.008)	(0.005)
Panel B: Secondary school participation								
Overall	-0.004	0.001	0.007	-0.010	0.000	0.017	0.004	-0.003
	(0.007)	(0.010)	(0.008)	(0.127)	(0.013)	(0.024)	(0.007)	(0.009)
Girls	0.232**	0.060	0.172	0.275*	0.131	-0.042	0.205*	0.186*
	(0.116)	(0.101)	(0.146)	(0.149)	(0.150)	(0.214)	(0.118)	(0.102)
Panel C: Performance for girls								
Passing rate	-1.131	-1.577	3.990	5.647	-0.469	-0.749	1.916	0.888
	(2.237)	(2.471)	(2.608)	(4.456)	(3.200)	(6.437)	(1.931)	(2.856)
Rate of distinction	-0.004	0.000	0.037**	-0.007	0.026	-0.028	0.011	-0.006
	(0.011)	(0.007)	(0.017)	(0.011)	(0.028)	(0.020)	(0.010)	(0.007)

Notes: Each cell reports the RD coefficient estimate at the cutoff based on Equation (10). Robust standard errors are shown in parentheses. Both running variables are re-centered so that the cutoff equals zero. Significance levels are denoted as follows: * $p < .1$, ** $p < .05$, *** $p < .01$.

This empirical association is also in line with the wider literature emphasizing that institutional quality conditions the productivity of educational resources. Wößmann (2003) argues that institutional arrangements are as important as resource levels for student achievement, while Kingdon and Teal (2010) and Suryadarma et al. (2006) show how weak accountability, and related governance failures can undermine learning in developing-country contexts. Similarly, Duflo et al. (2015) demonstrate that additional school resources need not improve outcomes in the absence of effective monitoring and implementation mechanisms. In Bangladesh, these institutional concerns are also well documented: school governing bodies often function only formally, monitoring is limited, and local political actors may influence school management and MPO continuation decisions (World Bank, 2013).

These considerations motivate a heterogeneity analysis along two dimensions. First, schools are divided into poorer and richer regions according to whether the local poverty headcount rate lies above or below the median. Second, schools are similarly stratified by the level of corruption in the education sector. The regression discontinuity design is then implemented separately across these strata in order to assess whether the causal effects of MPO funding differ systematically across regional socioeconomic and governance environments. This exercise is intended to examine whether the muted average effects

documented in the main analysis conceal meaningful heterogeneity across space.

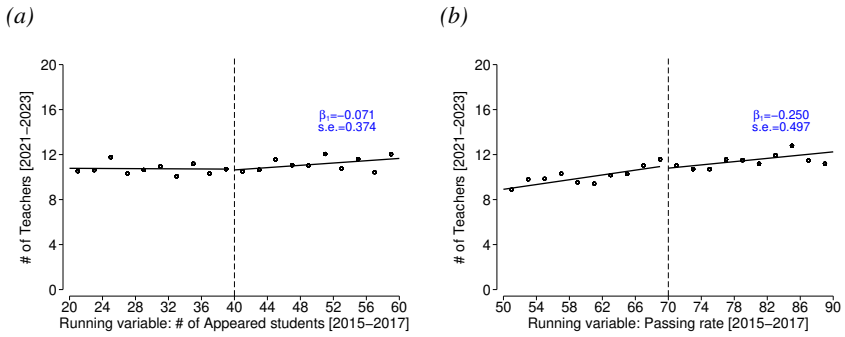
The result in table 4 does not point to meaningful heterogeneity. Although the descriptive evidence suggests that poverty and governance quality are strongly associated and plausibly relevant for school performance, the RD estimates do not reveal strong or consistent causal evidence that the effects of MPO funding on educational performance differ systematically across poverty or corruption strata. At most, the estimates point to isolated differences for selected outcomes, but these patterns are not sufficiently stable across specifications or running variables to support firm conclusions. Taken together, the findings suggest that regional socioeconomic and governance differentials are meaningfully associated with variation in school environments, but they do not translate into strong evidence of spatial heterogeneity in the causal impact of MPO funding on student performance.

Mechanism

The overall findings indicate that the teacher payroll subsidy program produced limited measurable impacts on school performance, except a modest increase in girls' secondary school participation. A closer look at the underlying mechanisms might help to understand the null effects.

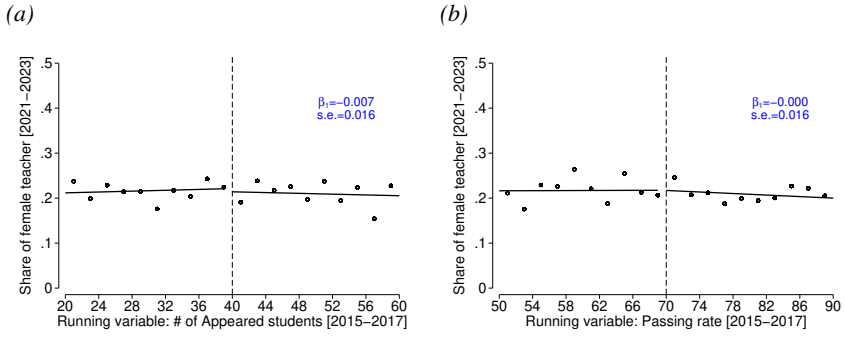
Stagnant school resources (# of teachers):
School resources are widely regarded as an important determinant of educational performance. A large empirical literature, in particular, highlights the role of class size in shaping student learning, with smaller classes often associated with more individualized attention, improved classroom management, and better instructional quality (e.g., Akerhielm, 1995; Krueger, 1999; Fredriksson et al., 2013; Behrman et al., 2008). One key determinant of class size is the number of teachers employed by the school. In this context, the MPO reform appears to have generated no change in this margin. Estimates for both the total number of teachers and the share of female teachers remain statistically insignificant across all specifications (figures 13 and 14), indicating no meaningful adjustment in either staffing levels or the gender composition of teachers following treatment. Given that total enrollment as well as gender composition in enrollment also remain largely unchanged, these results suggest that class sizes are unlikely to have changed materially in the post-reform period. The absence of staff expansion therefore offers one plausible explanation for the null effects on student achievement.

Figure 13. Outcome: # of total teachers



Notes: The figure plots the post-treatment average number of teachers against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure 14. Outcome: Share of female teacher



Notes: The figure plots the post-treatment average share of female teachers against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The post-treatment outcomes are averaged over 2021–2023. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

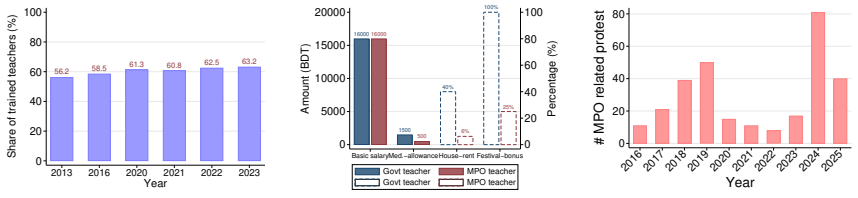
Moreover, while the subsidy may have eased schools' financial constraints by shifting payroll costs to the government, it remains unclear how the resulting savings, if any, were reallocated. The absence of observable changes in staffing suggests that these additional resources were not used to recruit more teachers and thereby improve instructional quality. However, staffing is unlikely to be the only binding constraint. In a study by the World Bank (2013), it was found that school-related factors account for 73 percent of the differences in students' performance, and many MPO schools face shortages of libraries, laboratories, multimedia facilities, and updated teaching materials, limiting opportunities

for interactive and practical learning and constraining their ability to convert additional financial support into improved educational outcomes.

Teacher quality, incentives, and accountability:
A further explanation for the null effects of MPO funding on student performance is that additional financial resources may have limited effects when teacher quality, effort, motivation, and accountability remain weak. Lack of school-level data related to these measures do not permit a direct test of the specific behavioral channels to evaluate our main results. However, the following discussion offers a set of logically consistent explanations with some descriptive and empirical evidence.

Figure 15. Quality, earnings disparity and dissatisfaction of school teachers in Bangladesh

(a) Share of trained teachers *(b)* Earnings differentials *(c)* MPO related protests



Notes: Figure 15a reports the share of secondary teachers in Bangladesh with the minimum required pedagogical qualifications, based on UNESCO SDG 4 statistics (UNESCO Institute for Statistics (2025)). The figure includes only years for which non-missing observations are available. Figure 15b presents a decomposition of compensation components for government and MPO teachers, distinguishing between monetary payments in Bangladeshi Taka and percentage-based benefits calculated as a share of basic salary. Figure 15c reports the annual frequency of MPO-related protest events derived from the GDELT database. These counts reflect media-reported protest events rather than necessarily unique protest occurrences, as multiple reports may cover the same demonstration.

A large literature identifies teacher quality as a central determinant of student achievement, with substantial variation in teacher effectiveness generating meaningful differences in both short-run academic outcomes and long-run adult outcomes (Rivkin et al., 2005; Rockoff, 2004; Chetty et al., 2014). One important dimension of teacher quality is pedagogical training. Figure 15a shows that roughly 40 percent of secondary school teachers in Bangladesh lack pedagogical training, and that this share has remained largely unchanged over the past decade. The situation is even more concerning in rural areas. According to the 2013 BANBEIS education statistics, only 24.5 percent of teachers in MPO-eligible rural schools had pedagogical training, compared with a national average exceeding 50 percent.

Beyond formal qualifications, however, the effectiveness of schooling also depends on whether teachers have the incentives and institutional support to translate their skills into classroom effort. Additional school inputs do not

automatically improve learning outcomes when other binding constraints remain reasonably weak. Mbiti et al. (2019) show that the returns to school resources depend critically on complementary factors such as teacher effort, while experimental evidence indicates that stronger monitoring and monetary incentives can improve teacher effort and raise student achievement (Muralidharan and Sundararaman, 2011; Duflo et al., 2012). In Bangladesh, however, the MPO reform expanded payroll support without materially changing the incentive environment in which teachers operate. Although the subsidy reduced schools' salary burden, it did not eliminate the compensation related discrimination faced by MPO teachers. As figure 15b shows, MPO and government teachers work under the same curriculum and examination system, yet MPO teachers receive substantially lower medical allowances, housing support, and festival bonuses, and they are not covered by pensions. These differences imply weaker long-term earnings security and may reduce motivation. The disadvantage is also broader: teachers in Bangladesh earn substantially less than their counterparts in neighboring lower-middle-income countries and in developed economies (appendix figure A1.2). Consistent with this, the World Bank (2013) also identifies salary disparities, weak recognition of performance, and limited accountability as important sources of low teacher motivation in Bangladesh. These economic grievances have translated into massive protests by MPO teachers in the recent years. Figure 15c documents the sustained MPO-related protest activity, suggesting persistent dissatisfaction over compensation and professional recognition.[17] Under such conditions, teachers may be more likely to supplement their income through private tutoring and coaching, potentially at the expense of classroom instruction (Deabnath, 2008); New Age (2024). More generally, Chaudhury et al. (2006) document that about $\frac{1}{6}^{th}$ of the school teachers in Bangladesh are absent from work on average, and high absenteeism is one important manifestation of low teacher effort (Mbiti et al., 2019).

One way to improve teacher effort is through stronger monitoring, performance-linked rewards, and more effective accountability mechanisms. Experimental evidence shows that monitoring can substantially reduce teacher absenteeism (Duflo et al., 2012), while linking incentives to performance can improve student achievement (Muralidharan and Sundararaman, 2011). In Bangladesh, however, the MPO system has struggled to function as a credible instrument for improving school outcomes, partly because governance weaknesses and political considerations have limited the enforcement of performance conditions. This problem is further compounded by weakly functioning school management committees, which may limit local account-

[17] It is important to note that the GDELT data capture media-reported events rather than strictly unique protest episodes, meaning that counts may include multiple reports of the same demonstration. Nevertheless, the pattern indicates sustained and, at times, intensifying contestation among teachers.

ability. Moreover, proposals to link teacher salaries to performance have previously faced strong resistance from MPO-recipient teachers and were ultimately abandoned under pressure (Asian Development Bank, 2015). Based on these descriptive evidence and institutional context one could argue that the MPO reform may have eased schools' payroll constraints without addressing the weaknesses in teacher quality, incentives, and accountability that shape classroom performance. If teacher effort remains low and the system does not effectively reward or monitor performance, additional payroll support alone is unlikely to generate measurable improvements in student learning. In this sense, the results of this paper are logically consistent with a setting in which payroll support was not accompanied by the reforms in teacher motivation, performance management, and accountability that may be necessary to convert public spending into measurable improvements in learning.

Conclusion

A longstanding question in the economics of education is whether expanding public financing for schools is sufficient to improve learning outcomes, particularly in low-income settings where resource constraints are severe and institutional challenges are widespread. This paper addresses that question through the first causal evaluation of Bangladesh's Monthly Pay Order (MPO) reform, a large-scale payroll subsidy for teachers in rural non-government secondary schools. Using nationally representative administrative data from 2012–2023 and a regression discontinuity design, the analysis examines whether sustained government financing of teacher salaries translates into improvements in student performance, school participation, and staffing.

The empirical results point to limited effects of the MPO reform on school performance. Across the main outcome measures, namely passing rates and the share of distinctions in the secondary school completion examination, there is no statistically significant evidence of improvement among MPO-recipient institutions. This suggests that the payroll subsidy did not translate into broad gains in student achievement. While the results point toward a modest expansion in girls' access to secondary schooling, the evidence is not strong enough to support a more definitive conclusion. I also examine heterogeneity by regional poverty and educational corruption, but find no meaningful differences in outcomes across these spatial dimensions. As a robustness check, results from the multi-score regression discontinuity design remain broadly consistent with the baseline sharp RDD estimates; i.e., no statistically meaningful treatment effects on performance outcomes, alongside only modest gains in girls' secondary school participation.

A plausible interpretation of these results is that the reform may have relaxed financial constraints without altering the core conditions that shape teaching and learning. The evidence shows no significant changes in teacher numbers or

staff composition, suggesting that class sizes remained largely unchanged and limiting potential gains in instructional quality. Although the payroll subsidy reduced schools' direct salary obligations, the lack of data makes it difficult to determine whether the resulting fiscal space was redirected toward investments that could improve teaching quality or learning conditions. At the same time, low teacher effort, weak monitoring, low performance thresholds, and the absence of credible accountability mechanisms likely reduced incentives for both teachers and administrators to exert additional effort. These institutional weaknesses are compounded by broader governance challenges, including corruption in the education sector, limited functionality of school management committees, and weak enforcement against underperformance (World Bank, 2013; Asian Development Bank, 2015). In addition, the persistent earnings gap between government teachers and MPO-supported teachers—reflected in lower allowances, the absence of pension entitlements, and recurrent protest activity—may have further undermined teacher motivation and classroom focus. These results are consistent with the previous literature (Duflo et al., 2015; Chin, 2005; Muralidharan and Sundararaman, 2011; Glewwe et al., 2010). Along with the causal evidence related to staffing, the descriptive evidence and institutional context provide a plausible explanation for why the reform failed to generate broad academic gains.

Overall, the paper yields two major insights. First, payroll subsidies may help sustaining the functioning of rural non-government secondary schools and may have modestly expanded girls' access to secondary education, despite lack of evidence on school outcomes. Second, unconditional salary support alone is unlikely to improve learning outcomes when schools lack the incentives, accountability, and institutional capacity needed to translate additional resources into better pedagogy. From a policy perspective, the MPO scheme may therefore represent an important yet incomplete intervention. To unlock the full potential of payroll subsidies, reforms must embed financial support within a broader framework of accountability, teacher development, and improved school-level governance. Strengthening oversight, addressing compensation inequities, improving instructional infrastructure, and better aligning incentives with student outcomes may be essential to achieving meaningful improvements

Declaration of Use of Generative AI

During the preparation of this manuscript, I used OpenAI's GPT-5.2 model to identify potential grammatical and typographical errors. All suggested corrections were carefully reviewed, and the manuscript was subsequently revised and finalized by me. I take full responsibility for the content of the article.

References

Aaronson, D. and Mazumder, B. (2011). The impact of rosenwald schools on black achievement. *Journal of Political Economy*, 119(5):821–888.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5):1369–1401.

Akerhielm, K. (1995). Does class size matter? *Economics of education Review*, 14(3):229–241.

Alderman, H., Kim, J., and Orazem, P. F. (2003). Design, evaluation, and sustainability of private schools for the poor: the pakistan urban and rural fellowship school experiments. *Economics of Education review*, 22(3):265–274.

Asian Development Bank (2015). Policy reform in bangladesh's secondary education (1993–2013): Tracing causal processes and examining the asian development bank's contribution. Topical paper, Asian Development Bank. Topical Paper, December 2015.

BANBEIS (2018). Bangladesh education statistics 2018.

Bank, W. (2024). Education statistics, human capital project (hcp). `https://humancapital.worldbank.org`. Accessed: September 14, 2025.

Bank, W., UNESCO Institute for Statistics (UIS), UN Educational, S., and (UNESCO), C. O. (2024). Education statistics, human capital project (hcp). `https://humancapital.worldbank.org`. Accessed: September 14, 2025.

Baron, E. J. (2022). School spending and student outcomes: Evidence from revenue limit elections in wisconsin. *American Economic Journal: Economic Policy*, 14(1):1–39.

Behrman, J. R., Ross, D., and Sabot, R. (2008). Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural pakistan. *Journal of Development Economics*, 85(1-2):94–104.

Bhatta, S. D., Genoni, M. E., Sharma, U., Khaltarkhuu, B. E., Maratou-Kolias, L., Asaduzzaman, T., et al. (2019). *Bangladesh education sector public expenditure review*. World Bank.

Cascio, E. U., Gordon, N., and Reber, S. (2013). Local responses to federal grants: Evidence from the introduction of title i in the south. *American Economic Journal: Economic Policy*, 5(3):126–159.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020a). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2020b). Analysis of regression-discontinuity designs with multiple cutoffs or multiple scores. *The Stata Journal*, 20(4):866–891.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., and Rogers, F. H. (2006). Missing in action: teacher and health worker absence in developing countries. *Journal of Economic perspectives*, 20(1):91–116.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, 104(9):2593–2632.

Chin, A. (2005). Can redistributing teachers across schools raise educational attainment? evidence from operation blackboard in india. *Journal of development Economics*, 78(2):384–405.

Deabnath, S. (2008). Bangladesh kids spend least time in schools among saarc states. *The Daily Star*. Front Page.

Duflo, E., Dupas, P., and Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from kenyan primary schools. *Journal of public Economics*, 123:92–110.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American economic review*, 102(4):1241–1278.

Dustmann, C., Rajah, N., and Van Soest, A. (2003). Class size, education, and wages. *The Economic Journal*, 113(485):F99–F120.

Fredriksson, P., Öckert, B., and Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly journal of economics*, 128(1):249–285.

Glewwe, P., Ilias, N., and Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, 2(3):205–227.

Goodman, S. F. and Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the new york city bonus program. *Journal of Labor Economics*, 31(2):409–420.

Hanushek, E. A. (2006). School resources. *Handbook of the Economics of Education*, 2:865–908.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The quarterly journal of economics*, 115(4):1239–1285.

Jackson, C. K. and Mackevicius, C. L. (2024). What impacts can we expect from school spending policy? evidence from evaluations in the united states. *American Economic Journal: Applied Economics*, 16(1):412–446.

Kaufmann, D., Kraay, A., and Zoido, P. (1999). Governance matters. *Available at SSRN 188568*.

Kingdon, G. and Teal, F. (2010). Teacher unions, teacher pay and student performance in india: A pupil fixed effects approach. *Journal of Development Economics*, 91(2):278–288.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8):2277–2304.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, 114(2):497–532.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, 10(2):1–26.

Lafortune, J. and Schönholzer, D. (2022). The impact of school facility investments on students and homeowners: Evidence from los angeles. *American Economic Journal: Applied Economics*, 14(3):254–289.

Leetaru, K. and Schrodt, P. A. (2013). Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, volume 2, pages 1–49. Citeseer.

Marshall, J. H. (2009). School quality and learning gains in rural guatemala. *Economics of Education Review*, 28(2):207–216.

Martorell, P., Stange, K., and McFarlin Jr, I. (2016). Investing in schools: capital spending, facility conditions, and student achievement. *Journal of Public Economics*, 140:13–29.

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., and Rajani, R. (2019). Inputs, incentives, and complementarities in education: Experimental evidence from tanzania. *The Quarterly journal of economics*, 134(3):1627–1673.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Menezes-Filho, N. and Pazello, E. (2007). Do teachers' wages matter for proficiency? evidence from a funding reform in brazil. *Economics of Education Review*, 26(6):660–672.

Ministry of Education, Bangladesh (2018). Organogram of private educational institutions and mpo policy 2018. Policy document, SHED, Ministry of Education, Bangladesh. Published in 2018.

Muralidharan, K. and Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from india. *Journal of political Economy*, 119(1):39–77.

New Age (2024). Govt should purge education of ills, put it back on track. *New Age*. Editorial.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World development*, 22(9):1325–1343.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *econometrica*, 73(2):417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American economic review*, 94(2):247–252.

Suryadarma, D., Suryahadi, A., Sumarto, S., and Rogers, F. H. (2006). Improving student performance in public primary schools in developing countries: Evidence from indonesia. *Education Economics*, 14(4):401–429.

UNESCO Institute for Statistics (2025). Sdg 4 country profile: Bangladesh.

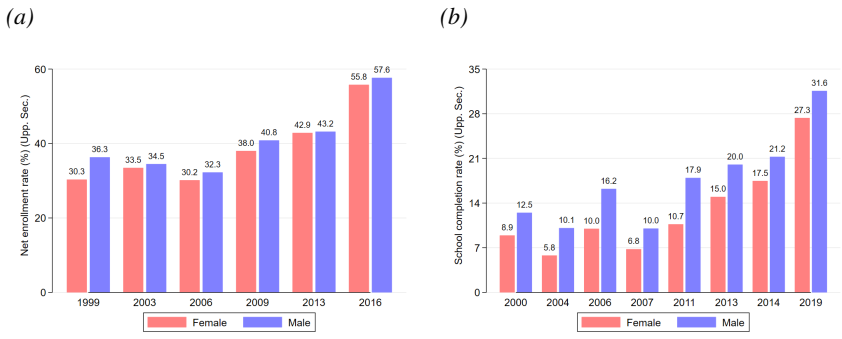
World Bank (2013). Bangladesh education sector review: Seeding fertile ground: Education that works for bangladesh. Sector review, The World Bank Office Dhaka, Washington, DC. Published in 2013.

Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford bulletin of economics and statistics*, 65(2):117–170.

Appendix A

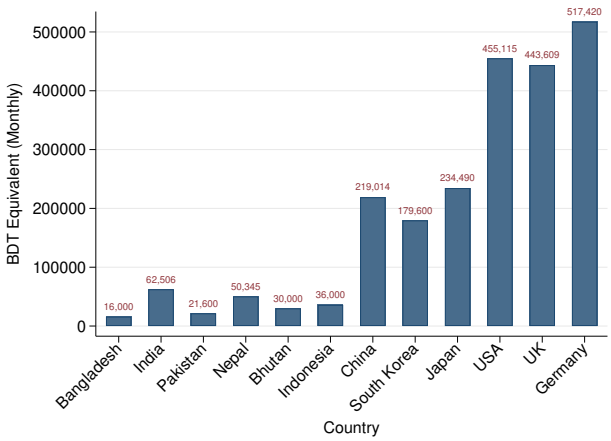
A1 Basic facts regarding secondary schools in Bangladesh

Figure A1.1. Secondary school statistics in Bangladesh (by gender)



Notes: Net Enrollment Rate (NER) is defined as the percentage of students in the official age group for secondary education who are enrolled in any level of education, calculated as the total number of students in the official age range enrolled in any level divided by the total population of the same age group, expressed as a percentage. School Completion Rate (SCR) is defined as the percentage of individuals who are 3–5 years older than the intended age for completing the last grade of secondary education and who have successfully completed that grade; for example, if the intended completion age is 18 years, the SCR reflects the percentage of individuals ages 21–23 who have completed the final grade (Source: World Bank, 2024).

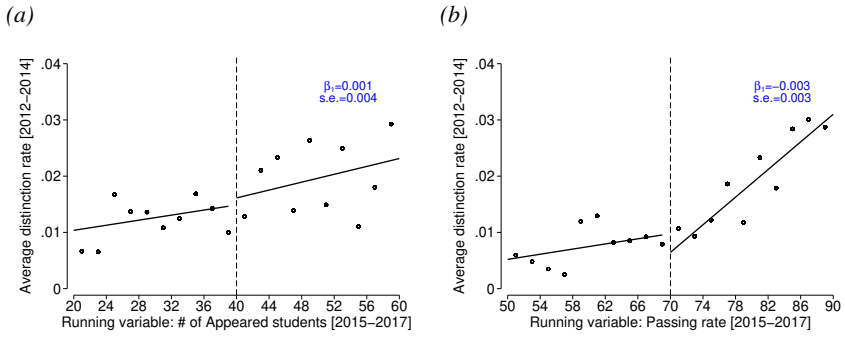
Figure A1.2. Monthly Teacher Salary in BDT Equivalent Across Countries



Notes: Monthly salary figures represent entry-level or starting teacher compensation converted to Bangladeshi Taka using exchange rates as of October 27, 2025. Data are compiled from national pay scales and publicly reported sources. Source: Md. Imtiaz Kabir Prottuy, "Bangladesh's Teachers Paid Least in Asia: Neighbours Earn Triple," *The Daily Campus*, 29 October 2025.

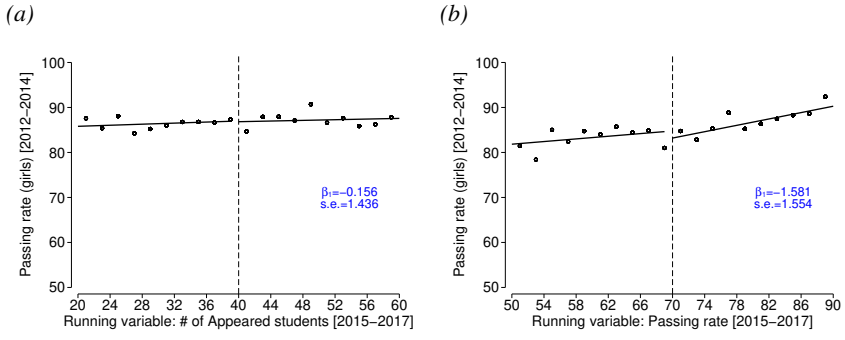
A2 Covariate balance

Figure A2.1. Covariate balance: Pre-treatment avg. distinction rate



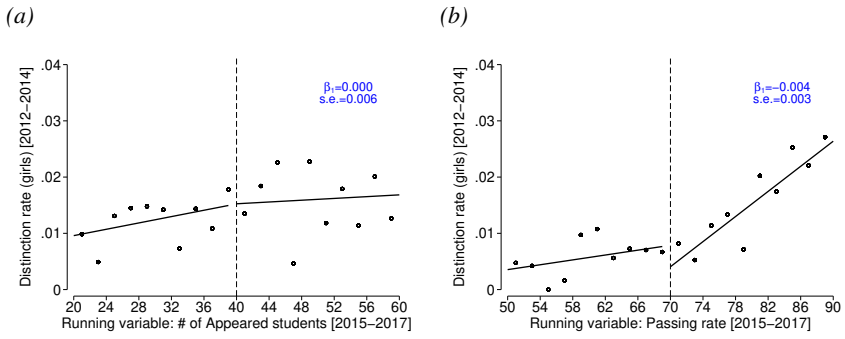
Notes: The figure plots the pre-treatment average distinction rate against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure A2.2. Covariate balance: Pre-determined average passing rate for girls



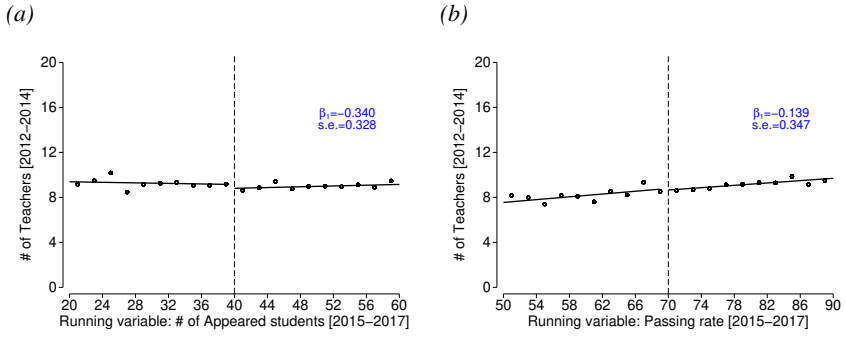
Notes: The figure plots the pre-treatment average passing rate for girls against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure A2.3. Covariate balance: Pre-treatment avg. distinction rate for girls



Notes: The figure plots the pre-treatment average distinction rate for girls against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure A2.4. Covariate balance: Pre-treatment # of total teachers



Notes: The figure plots the pre-treatment average number of teachers against (a) the average number of appeared students and (b) the average passing rate, with both running variables re-centered at zero to indicate the cutoff. The pre-treatment outcomes are averaged over 2012–2014. Each dot represents the midpoint of a bin (where the bin consists of two points), and the fitted regression lines are obtained from predicted values of the dependent variable using Equation (10). The estimated discontinuities at the threshold and their standard errors are reported in the figures. * $p < .1$, ** $p < .05$, *** $p < .01$.

A3 Bunching test with linear functional form

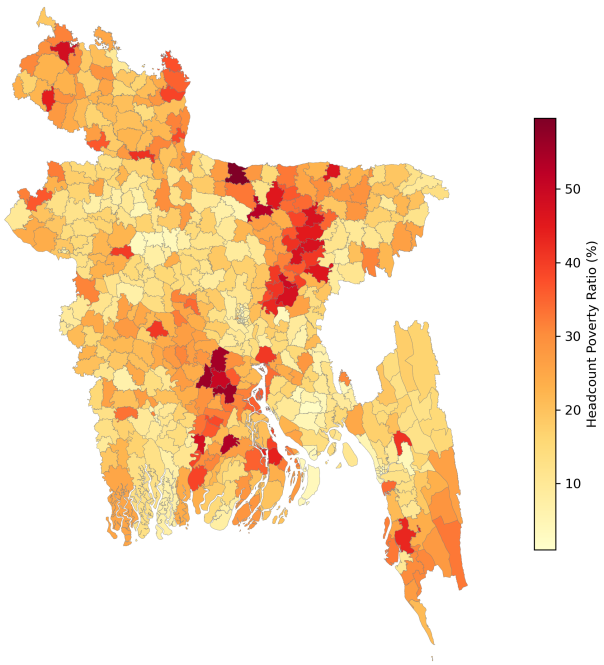
Table A3.1. *Bunching Test Results (Cattaneo, Jansson, and Ma, 2020)*

Running Variable	Test Statistic	p-value
Running Variable 1 (# of appeared students)	-0.461	0.645
Running Variable 2 (passing rate)	1.268	0.205

Notes: The table reports the test statistics from the manipulation test for bunching based on the density discontinuity test of Cattaneo et al. (2020a). The running variables are (1) the number of appeared students and (2) the passing rate. This specification uses a first-degree local polynomial to construct the density estimators. The test statistics and corresponding p-values are shown. Higher p-values indicate that the null hypothesis of no discontinuity in the density at the cutoff cannot be rejected, suggesting no evidence of manipulation or bunching around the threshold. Equal bandwidths on both sides of the cutoff were used for the test, and the results are robust to the data-driven bandwidth selectors proposed by Cattaneo et al. (2020a).

A4 Spatial variations in poverty level in Bangladesh

Figure A4.1. Local level poverty headcount



Notes: The figure presents the spatial distribution of poverty headcount ratios across local administrative regions of Bangladesh. The color gradient represents the intensity of poverty, with lighter shades indicating lower headcount ratios and darker red shades indicating higher headcount ratios. Regional boundary shapefiles are obtained from the Bangladesh Bureau of Statistics (BBS), and poverty headcount data are from the World Bank's Poverty Map of Bangladesh, 2022.

Appendix B

Multi-score Regression discontinuity design (MSRDD)

In a multi-dimensional regression discontinuity (RD) design, treatment assignment depends on multiple score variables or multiple cutoffs. The most common empirical case involves two running variables (Cattaneo et al., 2020b). In this study, treatment assignment is formally based on four running variables; however, we restrict the sample to schools that satisfy the thresholds on these less critical variables, such that variation in treatment status is effectively governed by the remaining two running variables. This design allows us to implement a multi-score regression discontinuity approach using the two binding assignment criteria. Unlike the one-dimensional RD case, where treatment is assigned at a single threshold, the multi-score RD design generates a boundary or surface across which treatment status changes. This results in a continuum of treatment effects defined along the boundary of the treatment region.

Each unit's score is represented by a vector, denoted by $\mathbf{R}_i = (R_{1i}, R_{2i})$, and treatment assignment is determined by a rule $D_i = D(\mathbf{R}_i)$. A common assignment rule is to require that both scores exceed a threshold, leading to the specification:

$$D(\mathbf{R}_i) = \mathbb{1}(R_{1i} > c_1) \cdot \mathbb{1}(R_{2i} > c_2),$$

where c_1 and c_2 are the respective cutoff points along each dimension. The potential outcomes are given by $Y_i(1)$ and $Y_i(0)$, and the treatment effect may be point-specific or pooled over the boundary. In the sharp RD framework, the point-specific treatment effect is denoted as:

$$\tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0) \mid \mathbf{R}_i = \mathbf{b}], \quad \mathbf{b} \in \mathscr{B},$$

where the treated and control regions are defined as:

$$\mathscr{A}_t = \{(r_1, r_2) : T(\mathbf{R}_i) = 1\}, \quad \mathscr{A}_c = \{(r_1, r_2) : T(\mathbf{R}_i) = 0\},$$

and the boundary between the treated and control areas is:

$$\mathscr{B} = \{(r_1, r_2) \in bd(\mathscr{A}_t) \cap bd(\mathscr{A}_c)\}, \quad bd(\mathscr{B}) \equiv cl(\mathscr{B}) \setminus int(\mathscr{B}).$$

In this study, MPO funding is granted if a school meets both of the following criteria: the three-year average number of students appearing in public examinations is at least 40, and the average passing rate is at least 70 percent. Denoting these variables as R_1 and R_2, respectively, the treatment boundary is expressed as:

$$\mathscr{B} = \{R_1 \geq 40, R_2 = 70\} \cup \{R_1 = 40, R_2 \geq 70\}.$$

Identification of point-specific treatment effects in a multi-score RD design closely parallels that of the single-score case. The treatment effect at a boundary point $\mathbf{b} \in \mathscr{B}$ is identified by the limit of the conditional expectation of the outcome variable from either side of the boundary:

$$\tau(\mathbf{b}) = \lim_{\mathbf{r}\to\mathbf{b},\mathbf{r}\in\mathscr{A}_t} \mathbb{E}[Y_i \mid \mathbf{R}_i = \mathbf{r}] - \lim_{\mathbf{r}\to\mathbf{b},\mathbf{r}\in\mathscr{A}_c} \mathbb{E}[Y_i \mid \mathbf{R}_i = \mathbf{r}].$$

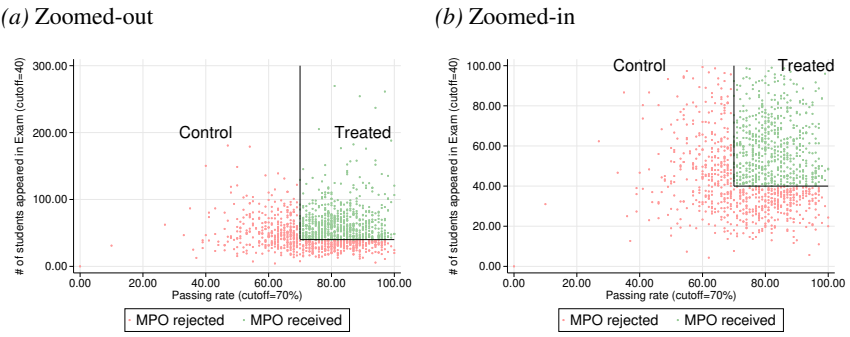
This formulation implies that the treatment effect at each boundary point $\mathbf{b}$ can be identified using observed bivariate regression functions estimated separately for treated and control units. Continuity of the running variables is crucial in this identification strategy. Although many empirical applications use discrete running variables that approximate continuity, such designs are still widely used. In the present case, the running variables are discrete as well. The multi-score RD setup naturally yields a set or curve of treatment effects $\tau(\mathbf{b})$, one for each boundary point, such as $\tau(40,70)$ or $\tau(40,80)$. A central identifying assumption in RD designs also applies here, which is the no-manipulation-with-precision assumption. This assumption implies that the running variables function like a randomizer around the cutoff: units close to the boundary cannot precisely manipulate their scores to select into treatment. If, however, the cutoff is not exogenous or units can deterministically influence their score to cross the threshold, this assumption fails. In such cases, the treatment assignment is no longer as good as random near the boundary. This assumption is testable and already discussed in the earlier sub-section.

Treatment assignment in multi-score RDD

Before presenting the estimates from the multi-score RDD, it is instructive to first visualize the two-dimensional space of the running variables in order to clearly define the treatment and control groups. Figure B1 illustrates the treatment assignment rule in the two-dimensional regression discontinuity (RD) framework. The vertical axis reports the three-year average number of students appearing in the examination (running variable 1), with a cutoff at 40, while the horizontal axis depicts the three-year average passing rate (running variable 2), with a cutoff at 70 percent. Schools that satisfy both eligibility thresholds-that is, those located in the upper-right quadrant of the diagram-constitute the treated group (green dots), as they receive MPO funding. In contrast, schools that fail to meet at least one of the criteria belong to the control group (pink dots), shown in the remaining quadrants. The discontinuity in treatment assignment therefore occurs along the L-shaped boundary defined by $\{R_1 \geq 40, R_2 = 70\} \cup \{R_1 = 40, R_2 \geq 70\}$. Local identification of causal effects relies on comparing outcomes of schools situated immediately on either side of this boundary, where, under the standard RD assumption of no precise manipulation of running variables, treatment status can be regarded as quasi-random. It is evident from the figure, however, that the support (number of observations) along the boundary is uneven, which raises concerns for

pooled estimates in this two-dimensional RD setting. Accordingly, the analysis emphasizes point-specific estimates at multiple locations along the boundary to assess whether results from the nonparametric multi-score RD design are consistent with the parametric single-score RD estimates presented earlier.

Figure B1. Treatment and control group in multi-score RDD setting

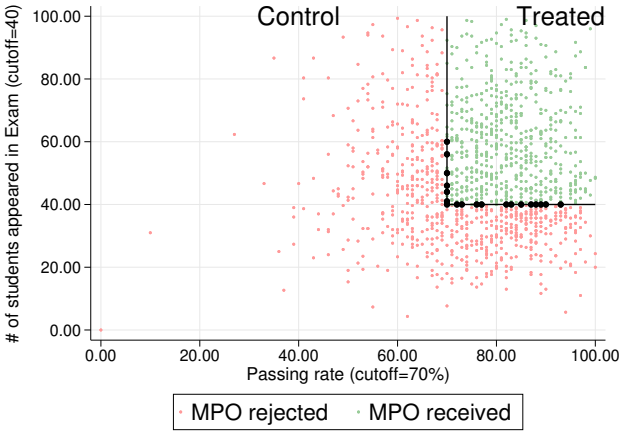


Notes: The figure illustrates the treatment and control regions in the multi-score regression discontinuity (RD) framework. Panel (a) shows the distribution of schools, while Panel (b) zooms in around the cutoff for clarity. The horizontal axis represents the passing rate, with a cutoff at 70 percent, and the vertical axis shows the number of students appearing in the exam (out of at least 40 students). Schools to the right of the vertical cutoff and above the horizontal cutoff are classified as treated (green), while those falling outside this region constitute the control group (red). This two-dimensional cutoff generates the treatment–control space for the multi-score RD design.

Based on visual inspection, a set of ordered points has been selected in the two-dimensional running variable space to serve as locations for point-specific estimates, as depicted in figure B2. These points were chosen at locations where the support on both sides of the treatment threshold appears relatively dense, thereby increasing the precision of the estimates and reducing noise. The vertical axis represents the average number of students appearing in examinations, which serves as a proxy for school size, while the horizontal axis captures the average pass rate, reflecting academic performance. Movement upward along the vertical axis corresponds to larger schools in terms of enrollment, whereas movement rightward along the horizontal axis corresponds to schools with stronger performance outcomes. Thus, each selected boundary point represents a distinct combination of school size and performance situated near the treatment threshold, enabling the analysis to disentangle heterogeneity

in treatment effects across different types of schools within the multi-score RD framework.

Figure B2. Boundary points for Multi-Score RD point estimation



Notes: This figure illustrates the treatment and control regions in the multi-score regression discontinuity (RD) framework, similar to figure B1. The set of solid black dots represents the selected points that have relatively higher support. These points are chosen as the cutoffs to conduct point-specific multi-score RD estimation proposed by Cattaneo et al. (2020b).

Findings from multi-score RDD

The multi-score regression discontinuity estimates, based on the methodology proposed by Cattaneo et al. (2020b), offer a nuanced picture of the causal effects of financial interventions on schools that were close to eligibility thresholds. In this paper, I deliberately avoid pooling estimates across the entire set of boundaries and instead focus on cutoff-specific robust bias-corrected estimates. This choice is both methodologically consistent with the challenges of uneven support around the boundary and also in line with the practical caveats highlighted in the existing literature. By estimating local linear regressions with MSE-optimal bandwidths, the estimates represent credible local average treatment effects at each cutoff point. The use of MSE-optimal bandwidth selectors ensures a balance between bias and variance, thereby yielding the most efficient estimates in finite samples. In addition, the regressions employ a triangular kernel, which assigns greater weight to observations closer to the cutoff while progressively down-weighting those further away. This weighting scheme is standard in the RD literature and particularly well-suited for local identification, as it sharpens inference by emphasizing the most comparable treatment and control units in the immediate neighborhood of the threshold.

The averaged results across the post-treatment years (2021–2023), in table B1, indicate that the intervention does not significantly affect core academic outcomes such as the overall passing rate, the passing rate for girls, or the share

of students achieving distinction, either for the full cohort or for girls specifically. This null effect is consistent across the main cutoff and several alternative boundaries. However, there are some weak significant and positive effects on measures of gender parity, particularly in the girl-to-boy overall enrollment ratio. As in the case of parametric estimation, the non-parametric estimates also show positive effects on girl enrollment over boy at the secondary level across several cut-off points at the treatment-control boundary, mostly for the better schools.[19] This effect suggests that the program might be effective in enhancing girls' enrollment at the targeted level. The interpretation of these effects requires understanding of the broader institutional context. The transition of schools from private to semi-public status through MPO eligibility appears to have created new incentives for parents to enroll their daughters, especially in institutions that had already achieved moderate to high pass rates. The sudden inflow of female students into these schools may partly reflect transfers from neighboring schools, suggesting that the intervention not only expanded access but also altered the competitive landscape of schooling options. It may also reflect the positive ex-ante expectation of the parents regarding their children's future academic performances. Finally, no systematic change is observed in the number of teachers employed.

[19]Better school is defined earlier as schools having higher passing rates the secondary public exam on average.

Table B1. *Multi-score RD estimates using average post-treatment outcomes*

Boundary cut-offs	Passing rate	Passing rate (girls)	Distinction share	Distinction share (girls)	Girls' secondary school participation	Number of teachers
(40,70)	1.281	4.821	0.011	0.009	-0.108	0.134
	(0.617)	(0.296)	(0.413)	(0.535)	(0.814)	(0.856)
(41,70)	1.472	5.119	0.011	0.013	-0.085	0.353
	(0.548)	(0.249)	(0.403)	(0.251)	(0.815)	(0.619)
(44,70)	1.046	4.790	0.012	0.006	-0.016	0.729
	(0.661)	(0.214)	(0.418)	(0.707)	(0.950)	(0.308)
(46,70)	0.049	1.744	0.017	0.007	0.030	0.822
	(0.983)	(0.673)	(0.241)	(0.678)	(0.863)	(0.300)
(50,70)	-0.806	-4.906	0.003	-0.013	0.215	0.936
	(0.697)	(0.469)	(0.785)	(0.442)	(0.237)	(0.372)
(56,70)	-0.451	3.052	-0.007	-0.010	0.208	1.243
	(0.833)	(0.592)	(0.596)	(0.581)	(0.464)	(0.237)
(60,70)	0.485	3.650	-0.005	-0.003	0.144	-0.290
	(0.858)	(0.547)	(0.696)	(0.833)	(0.511)	(0.767)
(40,72)	0.144	2.022	0.007	0.009	-0.004	-0.028
	(0.950)	(0.648)	(0.533)	(0.518)	(0.993)	(0.969)
(40,73)	-0.373	0.430	0.007	0.006	0.066	-0.059
	(0.872)	(0.920)	(0.464)	(0.630)	(0.834)	(0.935)
(40,76)	-0.901	-3.122	-0.004	-0.005	0.226	0.089
	(0.677)	(0.438)	(0.727)	(0.742)	(0.272)	(0.903)
(40,77)	-1.030	-4.219	-0.004	-0.008	0.252	0.082
	(0.616)	(0.269)	(0.659)	(0.578)	(0.159)	(0.917)
(40,82)	-0.389	-3.933	-0.007	-0.010	0.436**	0.003
	(0.842)	(0.154)	(0.519)	(0.474)	(0.020)	(0.997)
(40,83)	-0.023	-2.717	-0.002	-0.004	0.440**	-0.011
	(0.990)	(0.311)	(0.820)	(0.754)	(0.022)	(0.987)
(40,85)	0.903	-1.581	0.002	0.002	0.385**	-0.139
	(0.557)	(0.569)	(0.868)	(0.896)	(0.022)	(0.832)
(40,87)	1.588	-0.265	0.009	0.012	0.333**	-0.287
	(0.398)	(0.932)	(0.462)	(0.457)	(0.028)	(0.671)
(40,88)	1.850	0.794	0.012	0.017	0.312**	-0.387
	(0.334)	(0.803)	(0.320)	(0.307)	(0.049)	(0.563)
(40,89)	2.036	1.760	0.015	0.019	0.278*	-0.530
	(0.303)	(0.599)	(0.228)	(0.239)	(0.057)	(0.397)
(40,90)	2.201	2.626	0.016	0.021	0.241*	-0.587
	(0.279)	(0.407)	(0.209)	(0.223)	(0.093)	(0.380)

Notes: This table reports point-specific regression discontinuity estimates for average outcomes over 2021–2023, obtained using local linear regression with MSE-optimal bandwidths and the robust bias-corrected procedure of Cattaneo et al. (2020b). No additional covariates are included in the specification, so the estimates reflect the local treatment effect of crossing the respective cutoff without further adjustment. Boundary cut-offs are defined by two eligibility criteria: the number of examinees (first element in parentheses) and the passing rate (second element in parentheses). For example, the main cutoff $(40,70)$ requires at least 40 examinees and a passing rate of 70%. Varying the first element (e.g., $(41,70)$, $(44,70)$) captures changes in school size holding pass rate constant, while varying the second element (e.g., $(40,72)$, $(40,73)$) captures changes in pass rate holding school size constant. Outcome variables are expressed in their natural units: passing rates and distinction shares in percentage points; girls-to-boys enrollment ratios as proportions; and number of teachers in absolute counts. Reported values in parentheses are robust p-values from the bias-corrected procedure. Statistical significance is denoted by asterisks: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.10$.

Economic Studies

1987:1 Haraldson, Marty. To Care and To Cure. A linear programming approach to national health planning in developing countries. 98 pp.

1989:1 Chryssanthou, Nikos. The Portfolio Demand for the ECU. A Transaction Cost Approach. 42 pp.

1989:2 Hansson, Bengt. Construction of Swedish Capital Stocks, 1963-87. An Application of the Hulten-Wykoff Studies. 37 pp.

1989:3 Choe, Byung-Tae. Some Notes on Utility Functions Demand and Aggregation. 39 pp.

1989:4 Skedinger, Per. Studies of Wage and Employment Determination in the Swedish Wood Industry. 89 pp.

1990:1 Gustafson, Claes-Håkan. Inventory Investment in Manufacturing Firms. Theory and Evidence. 98 pp.

1990:2 Bantekas, Apostolos. The Demand for Male and Female Workers in Swedish Manufacturing. 56 pp.

1991:1 Lundholm, Michael. Compulsory Social Insurance. A Critical Review. 109 pp.

1992:1 Sundberg, Gun. The Demand for Health and Medical Care in Sweden. 58 pp.

1992:2 Gustavsson, Thomas. No Arbitrage Pricing and the Term Structure of Interest Rates. 47 pp.

1992:3 Elvander, Nils. Labour Market Relations in Sweden and Great Britain. A Comparative Study of Local Wage Formation in the Private Sector during the 1980s. 43 pp.

12 Dillén, Mats. Studies in Optimal Taxation, Stabilization, and Imperfect Competition. 1993. 143 pp.

13 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 1993. 303 pp.

14 Mellander, Erik. Measuring Productivity and Inefficiency Without Quantitative Output Data. 1993. 140 pp.

15 Ackum Agell. Susanne. Essays on Work and Pay. 1993. 116 pp.

16 Eriksson, Claes. Essays on Growth and Distribution. 1994. 129 pp.

17 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking and Finance. 2nd version, 1994. 313 pp.

18 Apel, Mikael. Essays on Taxation and Economic Behavior. 1994. 144 pp.

19 Dillén, Hans. Asset Prices in Open Monetary Economies. A Contingent Claims Approach. 1994. 100 pp.

20 Jansson, Per. Essays on Empirical Macroeconomics. 1994. 146 pp.

21 Banks, Ferdinand E.. A Modern Introduction to International Money, Banking, and Finance. 3rd version, 1995. 313 pp.

22 Dufwenberg, Martin. On Rationality and Belief Formation in Games. 1995. 93 pp.

23 Lindén, Johan. Job Search and Wage Bargaining. 1995. 127 pp.

24 Shahnazarian, Hovick. Three Essays on Corporate Taxation. 1996. 112 pp.

25 Svensson, Roger. Foreign Activities of Swedish Multinational Corporations. 1996. 166 pp.

26 Sundberg, Gun. Essays on Health Economics. 1996. 174 pp.

27 Sacklén, Hans. Essays on Empirical Models of Labor Supply. 1996. 168 pp.

28 Fredriksson, Peter. Education, Migration and Active Labor Market Policy. 1997. 106 pp.

29 Ekman, Erik. Household and Corporate Behaviour under Uncertainty. 1997. 160 pp.

30 Stoltz, Bo. Essays on Portfolio Behavior and Asset Pricing. 1997. 122 pp.

31 Dahlberg, Matz. Essays on Estimation Methods and Local Public Economics. 1997. 179 pp.

32 Kolm, Ann-Sofie. Taxation, Wage Formation, Unemployment and Welfare. 1997. 162 pp.

33 Boije, Robert. Capitalisation, Efficiency and the Demand for Local Public Services. 1997. 148 pp.

34 Hort, Katinka. On Price Formation and Quantity Adjustment in Swedish Housing Markets. 1997. 185 pp.

35 Lindström, Thomas. Studies in Empirical Macroeconomics. 1998. 113 pp.

36 Hemström, Maria. Salary Determination in Professional Labour Markets. 1998. 127 pp.

37 Forsling, Gunnar. Utilization of Tax Allowances and Corporate Borrowing. 1998. 96 pp.

38 Nydahl, Stefan. Essays on Stock Prices and Exchange Rates. 1998. 133 pp.

39 Bergström, Pål. Essays on Labour Economics and Econometrics. 1998. 163 pp.

40 Heiborn, Marie. Essays on Demographic Factors and Housing Markets. 1998. 138 pp.

41 Åsberg, Per. Four Essays in Housing Economics. 1998. 166 pp.

42 Hokkanen, Jyry. Interpreting Budget Deficits and Productivity Fluctuations. 1998. 146 pp.

43 Lunander, Anders. Bids and Values. 1999. 127 pp.

44 Eklöf, Matias. Studies in Empirical Microeconomics. 1999. 213 pp.

45 Johansson, Eva. Essays on Local Public Finance and Intergovernmental Grants. 1999. 156 pp.

46 Lundin, Douglas. Studies in Empirical Public Economics. 1999. 97 pp.

47 Hansen, Sten. Essays on Finance, Taxation and Corporate Investment. 1999. 140 pp.

48 Widmalm, Frida. Studies in Growth and Household Allocation. 2000. 100 pp.

49 Arslanogullari, Sebastian. Household Adjustment to Unemployment. 2000. 153 pp.

50 Lindberg, Sara. Studies in Credit Constraints and Economic Behavior. 2000. 135 pp.

51 Nordblom, Katarina. Essays on Fiscal Policy, Growth, and the Importance of Family Altruism. 2000. 105 pp.

52 Andersson, Björn. Growth, Saving, and Demography. 2000. 99 pp.

53 Åslund, Olof. Health, Immigration, and Settlement Policies. 2000. 224 pp.

54 Bali Swain, Ranjula. Demand, Segmentation and Rationing in the Rural Credit Markets of Puri. 2001. 160 pp.

55 Löfqvist, Richard. Tax Avoidance, Dividend Signaling and Shareholder Taxation in an Open Economy. 2001. 145 pp.

56 Vejsiu, Altin. Essays on Labor Market Dynamics. 2001. 209 pp.

57 Zetterström, Erik. Residential Mobility and Tenure Choice in the Swedish Housing Market. 2001. 125 pp.

58 Grahn, Sofia. Topics in Cooperative Game Theory. 2001. 106 pp.

59 Laséen, Stefan. Macroeconomic Fluctuations and Microeconomic Adjustments. Wages, Capital, and Labor Market Policy. 2001. 142 pp.

60 Arnek, Magnus. Empirical Essays on Procurement and Regulation. 2002. 155 pp.

61 Jordahl, Henrik. Essays on Voting Behavior, Labor Market Policy, and Taxation. 2002. 172 pp.

62 Lindhe, Tobias. Corporate Tax Integration and the Cost of Capital. 2002. 102 pp.

63 Hallberg, Daniel. Essays on Household Behavior and Time-Use. 2002. 170 pp.

64 Larsson, Laura. Evaluating Social Programs: Active Labor Market Policies and Social Insurance. 2002. 126 pp.

65 Bergvall, Anders. Essays on Exchange Rates and Macroeconomic Stability. 2002. 122 pp.

66 Nordström Skans, Oskar. Labour Market Effects of Working Time Reductions and Demographic Changes. 2002. 118 pp.

67 Jansson, Joakim. Empirical Studies in Corporate Finance, Taxation and Investment. 2002. 132 pp.

68 Carlsson, Mikael. Macroeconomic Fluctuations and Firm Dynamics: Technology, Production and Capital Formation. 2002. 149 pp.

69 Eriksson, Stefan. The Persistence of Unemployment: Does Competition between Employed and Unemployed Job Applicants Matter? 2002. 154 pp.

70 Huitfeldt, Henrik. Labour Market Behaviour in a Transition Economy: The Czech Experience. 2003. 110 pp.

71 Johnsson, Richard. Transport Tax Policy Simulations and Satellite Accounting within a CGE Framework. 2003. 84 pp.

72 Öberg, Ann. Essays on Capital Income Taxation in the Corporate and Housing Sectors. 2003. 183 pp.

73 Andersson, Fredrik. Causes and Labor Market Consequences of Producer Heterogeneity. 2003. 197 pp.

74 Engström, Per. Optimal Taxation in Search Equilibrium. 2003. 127 pp.

75 Lundin, Magnus. The Dynamic Behavior of Prices and Investment: Financial Constraints and Customer Markets. 2003. 125 pp.

76 Ekström, Erika. Essays on Inequality and Education. 2003. 166 pp.

77 Barot, Bharat. Empirical Studies in Consumption, House Prices and the Accuracy of European Growth and Inflation Forecasts. 2003. 137 pp.

78 Österholm, Pär. Time Series and Macroeconomics: Studies in Demography and Monetary Policy. 2004. 116 pp.

79 Bruér, Mattias. Empirical Studies in Demography and Macroeconomics. 2004. 113 pp.

80 Gustavsson, Magnus. Empirical Essays on Earnings Inequality. 2004. 154 pp.

81 Toll, Stefan. Studies in Mortgage Pricing and Finance Theory. 2004. 100 pp.

82 Hesselius, Patrik. Sickness Absence and Labour Market Outcomes. 2004. 109 pp.

83 Häkkinen, Iida. Essays on School Resources, Academic Achievement and Student Employment. 2004. 123 pp.

84 Armelius, Hanna. Distributional Side Effects of Tax Policies: An Analysis of Tax Avoidance and Congestion Tolls. 2004. 96 pp.

85 Ahlin, Åsa. Compulsory Schooling in a Decentralized Setting: Studies of the Swedish Case. 2004. 148 pp.

86 Heldt, Tobias. Sustainable Nature Tourism and the Nature of Tourists' Cooperative Behavior: Recreation Conflicts, Conditional Cooperation and the Public Good Problem. 2005. 148 pp.

87 Holmberg, Pär. Modelling Bidding Behaviour in Electricity Auctions: Supply Function Equilibria with Uncertain Demand and Capacity Constraints. 2005. 43 pp.

88 Welz, Peter. Quantitative new Keynesian macroeconomics and monetary policy 2005. 128 pp.

89 Ågren, Hanna. Essays on Political Representation, Electoral Accountability and Strategic Interactions. 2005. 147 pp.

90 Budh, Erika. Essays on environmental economics. 2005. 115 pp.

91 Chen, Jie. Empirical Essays on Housing Allowances, Housing Wealth and Aggregate Consumption. 2005. 192 pp.

92 Angelov, Nikolay. Essays on Unit-Root Testing and on Discrete-Response Modelling of Firm Mergers. 2006. 127 pp.

93 Savvidou, Eleni. Technology, Human Capital and Labor Demand. 2006. 151 pp.

94 Lindvall, Lars. Public Expenditures and Youth Crime. 2006. 112 pp.

95 Söderström, Martin. Evaluating Institutional Changes in Education and Wage Policy. 2006. 131 pp.

96 Lagerström, Jonas. Discrimination, Sickness Absence, and Labor Market Policy. 2006. 105 pp.

97 Johansson, Kerstin. Empirical essays on labor-force participation, matching, and trade. 2006. 168 pp.

98 Ågren, Martin. Essays on Prospect Theory and the Statistical Modeling of Financial Returns. 2006. 105 pp.

99 Nahum, Ruth-Aïda. Studies on the Determinants and Effects of Health, Inequality and Labour Supply: Micro and Macro Evidence. 2006. 153 pp.

100 Žamac, Jovan. Education, Pensions, and Demography. 2007. 105 pp.

101 Post, Erik. Macroeconomic Uncertainty and Exchange Rate Policy. 2007. 129 pp.

102 Nordberg, Mikael. Allies Yet Rivals: Input Joint Ventures and Their Competitive Effects. 2007. 122 pp.

103 Johansson, Fredrik. Essays on Measurement Error and Nonresponse. 2007. 130 pp.

104 Haraldsson, Mattias. Essays on Transport Economics. 2007. 104 pp.

105 Edmark, Karin. Strategic Interactions among Swedish Local Governments. 2007. 141 pp.

106 Oreland, Carl. Family Control in Swedish Public Companies. Implications for Firm Performance, Dividends and CEO Cash Compensation. 2007. 121 pp.

107 Andersson, Christian. Teachers and Student Outcomes: Evidence using Swedish Data. 2007. 154 pp.

108 Kjellberg, David. Expectations, Uncertainty, and Monetary Policy. 2007. 132 pp.

109 Nykvist, Jenny. Self-employment Entry and Survival - Evidence from Sweden. 2008. 94 pp.

110 Selin, Håkan. Four Empirical Essays on Responses to Income Taxation. 2008. 133 pp.

111 Lindahl, Erica. Empirical studies of public policies within the primary school and the sickness insurance. 2008. 143 pp.

112 Liang, Che-Yuan. Essays in Political Economics and Public Finance. 2008. 125 pp.

113 Elinder, Mikael. Essays on Economic Voting, Cognitive Dissonance, and Trust. 2008. 120 pp.

114 Grönqvist, Hans. Essays in Labor and Demographic Economics. 2009. 120 pp.

115 Bengtsson, Niklas. Essays in Development and Labor Economics. 2009. 93 pp.

116 Vikström, Johan. Incentives and Norms in Social Insurance: Applications, Identification and Inference. 2009. 205 pp.

117 Liu, Qian. Essays on Labor Economics: Education, Employment, and Gender. 2009. 133 pp.

118 Glans, Erik. Pension reforms and retirement behaviour. 2009. 126 pp.

119 Douhan, Robin. Development, Education and Entrepreneurship. 2009.

120 Nilsson, Peter. Essays on Social Interactions and the Long-term Effects of Early-life Conditions. 2009. 180 pp.

121 Johansson, Elly-Ann. Essays on schooling, gender, and parental leave. 2010. 131 pp.

122 Hall, Caroline. Empirical Essays on Education and Social Insurance Policies. 2010. 147 pp.

123 Enström-Öst, Cecilia. Housing policy and family formation. 2010. 98 pp.

124 Winstrand, Jakob. Essays on Valuation of Environmental Attributes. 2010. 96 pp.

125 Söderberg, Johan. Price Setting, Inflation Dynamics, and Monetary Policy. 2010. 102 pp.

126 Rickne, Johanna. Essays in Development, Institutions and Gender. 2011. 138 pp.

127 Hensvik, Lena. The effects of markets, managers and peers on worker outcomes. 2011. 179 pp.

128 Lundqvist, Heléne. Empirical Essays in Political and Public. 2011. 157 pp.

129 Bastani, Spencer. Essays on the Economics of Income Taxation. 2012. 257 pp.

130 Corbo, Vesna. Monetary Policy, Trade Dynamics, and Labor Markets in Open Economies. 2012. 262 pp.

131 Nordin, Mattias. Information, Voting Behavior and Electoral Accountability. 2012. 187 pp.

132 Vikman, Ulrika. Benefits or Work? Social Programs and Labor Supply. 2013. 161 pp.

133 Ek, Susanne. Essays on unemployment insurance design. 2013. 136 pp.

134 Österholm, Göran. Essays on Managerial Compensation. 2013. 143 pp.

135 Adermon, Adrian. Essays on the transmission of human capital and the impact of technological change. 2013. 138 pp.

136 Kolsrud, Jonas. Insuring Against Unemployment 2013. 140 pp.

137 Hanspers, Kajsa. Essays on Welfare Dependency and the Privatization of Welfare Services. 2013. 208 pp.

138 Persson, Anna. Activation Programs, Benefit Take-Up, and Labor Market Attachment. 2013. 164 pp.

139 Engdahl, Mattias. International Mobility and the Labor Market. 2013. 216 pp.

140 Krzysztof Karbownik. Essays in education and family economics. 2013. 182 pp.

141 Oscar Erixson. Economic Decisions and Social Norms in Life and Death Situations. 2013. 183 pp.

142 Pia Fromlet. Essays on Inflation Targeting and Export Price Dynamics. 2013. 145 pp.

143 Daniel Avdic. Microeconometric Analyses of Individual Behavior in Public Welfare Systems. Applications in Health and Education Economics. 2014. 176 pp.

144 Arizo Karimi. Impacts of Policies, Peers and Parenthood on Labor Market Outcomes. 2014. 221 pp.

145 Karolina Stadin. Employment Dynamics. 2014. 134 pp.

146 Haishan Yu. Essays on Environmental and Energy Economics. 132 pp.

147 Martin Nilsson. Essays on Health Shocks and Social Insurance. 139 pp.

148 Tove Eliasson. Empirical Essays on Wage Setting and Immigrant Labor Market Opportunities. 2014. 144 pp.

149 Erik Spector. Financial Frictions and Firm Dynamics. 2014. 129 pp.

150 Michihito Ando. Essays on the Evaluation of Public Policies. 2015. 193 pp.

151 Selva Bahar Baziki. Firms, International Competition, and the Labor Market. 2015. 183 pp.

152 Fredrik Sävje. What would have happened? Four essays investigating causality. 2015. 229 pp.

153 Ina Blind. Essays on Urban Economics. 2015. 197 pp.

154 Jonas Poulsen. Essays on Development and Politics in Sub-Saharan Africa. 2015. 240 pp.

155 Lovisa Persson. Essays on Politics, Fiscal Institutions, and Public Finance. 2015. 137 pp.

156 Gabriella Chirico Willstedt. Demand, Competition and Redistribution in Swedish Dental Care. 2015. 119 pp.

157 Yuwei Zhao de Gosson de Varennes. Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes. 2016. 148 pp.

158 Johannes Hagen. Essays on Pensions, Retirement and Tax Evasion. 2016. 195 pp.

159 Rachatar Nilavongse. Housing, Banking and the Macro Economy. 2016. 156 pp.

160 Linna Martén. Essays on Politics, Law, and Economics. 2016. 150 pp.

161 Olof Rosenqvist. Essays on Determinants of Individual Performance and Labor Market Outcomes. 2016. 151 pp.

162 Linuz Aggeborn. Essays on Politics and Health Economics. 2016. 203 pp.

163 Glenn Mickelsson. DSGE Model Estimation and Labor Market Dynamics. 2016. 166 pp.

164 Sebastian Axbard. Crime, Corruption and Development. 2016. 150 pp.

165 Mattias Öhman. Essays on Cognitive Development and Medical Care. 2016. 181 pp.

166 Jon Frank. Essays on Corporate Finance and Asset Pricing. 2017. 160 pp.

167 Ylva Moberg. Gender, Incentives, and the Division of Labor. 2017. 220 pp.

168 Sebastian Escobar. Essays on inheritance, small businesses and energy consumption. 2017. 194 pp.

169 Evelina Björkegren. Family, Neighborhoods, and Health. 2017. 226 pp.

170 Jenny Jans. Causes and Consequences of Early-life Conditions. Alcohol, Pollution and Parental Leave Policies. 2017. 209 pp.

171 Josefine Andersson. Insurances against job loss and disability. Private and public interventions and their effects on job search and labor supply. 2017. 175 pp.

172 Jacob Lundberg. Essays on Income Taxation and Wealth Inequality. 2017. 173 pp.

173 Anna Norén. Caring, Sharing, and Childbearing. Essays on Labor Supply, Infant Health, and Family Policies. 2017. 206 pp.

174 Irina Andone. Exchange Rates, Exports, Inflation, and International Monetary Cooperation. 2018. 174 pp.

175 Henrik Andersson. Immigration and the Neighborhood. Essays on the Causes and Consequences of International Migration. 2018. 181 pp.

176 Aino-Maija Aalto. Incentives and Inequalities in Family and Working Life. 2018. 131 pp.

177 Gunnar Brandén. Understanding Intergenerational Mobility. Inequality, Student Aid and Nature-Nurture Interactions. 2018. 125 pp.

178 Mohammad H. Sepahvand. Essays on Risk Attitudes in Sub-Saharan Africa. 2019. 215 pp.

179 Mathias von Buxhoeveden. Partial and General Equilibrium Effects of Unemployment Insurance. Identification, Estimation and Inference. 2019. 89 pp.

180 Stefano Lombardi. Essays on Event History Analysis and the Effects of Social Programs on Individuals and Firms. 2019. 150 pp.

181 Arnaldur Stefansson. Essays in Public Finance and Behavioral Economics. 2019. 191 pp.

182 Cristina Bratu. Immigration: Policies, Mobility and Integration. 2019. 173 pp.

183 Tamás Vasi. Banks, Shocks and Monetary Policy. 2020. 148 pp.
184 Jonas Cederlöf. Job Loss: Consequences and Labor Market Policy. 2020. 213 pp.

185 Dmytro Stoyko. Expectations, Financial Markets and Monetary Policy. 2020. 153 pp.

186 Paula Roth. Essays on Inequality, Insolvency and Innovation. 2020. 191 pp.

187 Fredrik Hansson. Consequences of Poor Housing, Essays on Urban and Health Economics. 2020. 143 pp.

188 Maria Olsson. Essays on Macroeconomics: Wage Rigidity and Aggregate Fluctuations. 2020. 130 pp.

189 Dagmar Müller. Social Networks and the School-to-Work Transition. 2020. 146 pp.

190 Maria Sandström. Essays on Savings and Intangible Capital. 2020. 129 pp.

191. Anna Thoresson. Wages and Their Impact on Individuals, Households and Firms. 2020. 220 pp.

192. Jonas Klarin. Empirical Essays in Public and Political Economics. 2020. 129 pp.

193. André Reslow. Electoral Incentives and Information Content in Macroeconomic Forecasts. 2021. 184 pp.

194. Davide Cipullo. Political Careers, Government Stability, and Electoral Cycles. 2021. 308 pp.

195. Olle Hammar. The Mystery of Inequality: Essays on Culture, Development, and Distributions. 2021. 210 pp.

196. J. Lucas Tilley. Inputs and Incentives in Education. 2021. 184 pp.

197. Sebastian Jävervall. Corruption, Distortions and Development. 2021. 215 pp.

198. Vivika Halapuu. Upper Secondary Education: Access, Choices and Graduation. 2021. 141 pp.

199. Charlotte Paulie. Essays on the Distribution of Production, Prices and Wealth. 2021. 141 pp.

200. Kerstin Westergren. Essays on Inflation Expectations, Monetary Policy and Tax Reform. 2021. 124 pp.

201. Melinda Süveg. Finance, Shocks, Competition and Price Setting. 2021. 137 pp.

202. Adrian Poignant. Gold, Coal and Iron. Essays on Industrialization and Economic Development. 2022. 214 pp.

203. Daniel Bougt. A Sequence of Essays on Sequences of Auctions. 2022. 188 pp.

204. Lillit Ottosson. From Welfare to Work. Financial Incentives, Active Labor Market Policies, and Integration Programs. 2022. 219 pp.

205. Yaroslav Yakymovych. Workers and Occupations in a Changing Labour Market. The Heterogeneous Effects of Mass Layoffs and Social Safety Nets. 2022. 212 pp.

206. Raoul van Maarseveen. Urbanization and Education. The Effect of Childhood Urban Residency on Educational Attainment. 2022. 210 pp.

207. Sofia Hernnäs. Automation and the Consequences of Occupational Decline. 2022. 229 pp.

208. Simon Ek, Structural Change, Match Quality, and Integration in the Labor Market. 2023. 176 pp.

209. Maximiliano Sosa Andres, A Risky and Polarized World: Essays on Uncertainty, Ideology and Foreign Policy. 2023. 226 pp.

210. Alice Hallman, Hypocrites, Devil's Advocates, and Bandwagoneers. Essays on Costly Signaling. 2023. 176 pp.

211. Edvin Hertegård, Essays on Families, Health Policy, and the Determinants of Children's Long-Term Outcomes. 2023. 236 pp.

212. Josefin Videnord, Trade, Innovation, and Gender. 2023. 156 pp.

213. Ao Fei, Essays on Taxation, Externalities, and Poverty Traps. 2023. 127 pp.

214. Anton Sundberg, Essays in Labor Economics. Parenthood, Immigration, and Education. 2024. 209 pp.

215. Gabriella Kindström, Urban Dynamics and Contemporary Challenges. Essays on Housing and Neighborhood Amenities. 2024. 161 pp.

216. Akib Khan, Essays in Human Capital and Immigration. 2024. 308 pp.

217. Erika Forsberg, Labor-market inequality. Essays on the roles of families, firms, location, and criminal records. 2024. 218 pp.

218. Qingyan Peng, Technology, Nonstandard Jobs, Occupations, and Firm Performance. 2024. 110 pp.

219. Hiep Nguyen, Essays on Environmental and Monetary Policy. 2024. 171 pp.

220. Markus Ridder, Essays on Polarization and Inequality. 2024. 185 pp.

221. Zunyuan Zheng, Cancer Experience and Charitable Giving. 2024. 176 pp.

222. Malin Backman, Essays on Women in the Labor Market. Technology, Inequality, and the Future of Work. 2025. 182 pp.

223. Hanfeng Chen, Rethinking Monetary Policy Transmission. Nonbank Finance, Central Bank Digital Currencies, and the Role of Bank Market Power. 2025. 169 pp.

224. Sundberg, Elin, Essays in Labor Economics. Gender, Careers, and Family Decisions. 2025. 276 pp.

225. Hussain, Amin, Optimizing Healthcare Payments. Regulation, Payment Schemes and Value-Adjustment. 2025. 222 pp.

226. Gülümser, Dogan, Wage Setting and the Gender Wage Gap. Essays on Inequality in Imperfect Labor Markets. 2025. 227 pp.

227. Liu, Tianze, Essays on the Economics of Specialized Skills. 2025. 125 pp.

228. Filippin, Maria Elena, The Macroeconomics of Digital Money. Household Adoption, Bank Intermediation, and Monetary Policy. 2025. 153 pp.

229. Karlsson, Kristina, Essays on Labor Economics, Household Savings, and Pension Reforms. 2025. 177 pp.

230. Czarnota, Alexander, On Monetary Policy, Prices, and Uncertainty. 2025. 145 pp.

231. Armerin, Fredrik, Studies on Uncertainties in Economics. 2025. 126 pp.

232. Koslyk, Caio, Essays on Quantitative Macroeconomics. News, Unemployment, and Fiscal Multipliers. 2025. 136 pp.

233. Gill, Adam, Beyond the Office Chair. Working from Home, Labor Market Access, and the Spatial Distribution of Firms. 2026. 254 pp.

234. Kwok, Tsz Chun, Essays on the Economics of Family, Fertility, and Inequality. 2026. 164 pp.

235. Rambjer, Lovisa, Algorithms, Incentives, and Information. Essays on Labor Markets and Public Service Provision. 2026. 199 pp.

236. Hossain, Mohammad Zariab, Networks, Norms, and Non-market Subsidy. Three essays on Public Policy. 2026. 140 pp.