The Indian Labor Market through the Lens of Public Sector Recruitment

Insights from the Tamil Nadu Public Service Commission to Inform Labor Market Policy and Improve Recruitment Practice

Kunal Mangal

March 13, 2023

Acknowledgments

This report is the culmination of an effort that began in 2018, and which would not have been possible without the support, effort, and faith of many, many people.

First, none of this would have been possible without the vision and support of officers within the Tamil Nadu government. S. Nagarajan and K. Nanthakumar believed in the potential of this project early on, and provided the right environment for it to take root. The able leadership of R. Sudhan, P. Uma Maheshwari, and Kiran Gurrala helped ensure that the seed of the idea was able to grow and develop.

Second, I am grateful to the staff at the Tamil Nadu Public Commission, whose extensive cooperation made it possible to learn, obtain the necessary data, and acquire the context needed to interpret the findings.

Third, I am grateful to the School Education department for hosting me during my many visits to Tamil Nadu. In particular, Archana Hari helped me find a second home in Tamil Nadu, and I will never forget her friendship and warmth.

Fourth, there are many people to thank for their help in the research process itself. Niharika Singh has been a brilliant collaborator. The Candidate Survey and Chapter 4 reflect joint work with her, and I am grateful for her companionship over the course of this research journey. Sivaperuman M., Deepa E., Aswinah Zeenath, and Vasudevan provided substantial support during the early days of piloting. S. Manohar provided excellent assistance. Karthik Srinivasan and Dr. Preethi Srinivasan provided substantial help in distributing the Candidate Survey. I am grateful to Azim Premji University—in particular Amit Basole—for hosting me while I undertook this project. Apurva Bamezai and Manchit Jasrotia provided valuable feedback on early drafts.

Finally, and most importantly, I am grateful to the many focus group participants and the thousands of survey respondents for sharing their experiences with us, even if that meant opening up about sometimes quite difficult parts of their life.

I gratefully acknowledge funding from the Weiss Fund for Research in Development Economics at the University of Chicago. The Candidate Survey was supported through funding from INCITE @ Columbia University. The data collection protocols for the Candidate Survey and its followup were reviewed and approved by the Institutional Review Board at Azim Premji University.

The views expressed in this report are those of the author alone, and do not reflect the views of the Tamil Nadu government, the Tamil Nadu Public Service Commission, any of the author's affiliations, or any of the funders. Any errors remain the author's alone.

About the Author

Kunal Mangal holds a PhD in Public Policy from Harvard University. He is currently a Visiting Fellow at the Centre for Sustainable Employment at Azim Premji University. His website is: `kmangal.github.io`.

Table of Contents

Version: v1.01

List of Figures

List of Tables

Introduction

Objectives of this Report

In India, as in many countries around the world, most government jobs are allocated through a system of merit-based exams. Over the past few decades, these exams have become incredibly competitive, at times receiving over 1000 applications for each vacancy.

Against a backdrop of rising educational attainment, high aspirations, disappointment with private sector opportunities, and a deep unmet need for income security, it is understandable why the demand for public sector employment opportunities is so high.

Yet despite the large footprint that public sector recruitments have in our social, economic and political life, many basic questions about them remain shrouded in mystery. Who applies? Why do they apply? Are these intense levels of competition helpful, or do they make people worse off? Why are people willing to invest so much in exam preparation? Why are people willing to gamble on such low odds of getting selected? What is the best way of structuring recruitments for both candidates and the government?

Our lack of understanding limits our ability to formulate sound labor market policy. As we will see, a large share of college graduates participate in public sector recruitment exams, and candidates for these exams make up a disproportionate share of the overall unemployed population. How can we improve employment outcomes if we do not understand who the unemployed are and how they are currently investing their time?

The main reason for the holes in our understanding is a lack of data. To date, neither private nor public household surveys include questions on whether individuals are preparing for competitive exams; and recruitment agencies have historically been cloistered institutions, understandably concerned about protecting the integrity of the recruitment process. As a result, the crores of candidates preparing for competitive exams around the country remain largely invisible in data, and by extension in policy.

This report attempts to shine a light on this dark corner of the labor market. To do so, I use several new sources of data. First, I draw on administrative data that allows us to observe the whole recruitment process for the entire universe of applicants. This is the first time such data has been made available in the Indian context. Second, I use data from a large-scale survey of over 3,000 candidates, which provides information about their investments in exam preparation, their access to resources, their constraints, and their beliefs. Third, my research collaborators and I conducted interviews and focus groups with candidates to better understand them in their own words. These rich data sources provide us with new insights into the economic and social life of candidates preparing for competitive exams.

The goal of this report is to demonstrate how both labor market and recruitment policy can be informed by a better understanding of candidate application behavior. This understanding can, in turn, help us tackle some of the key challenges in the modern Indian labor market—high levels of educated unemployment, a lack of skill development, low levels of female labor force participation, and more.

Still, I do not expect to have answers to all the important questions yet. There are important limitations to what I can say with confidence (which I will highlight throughout the report). My hope is that this report encourages more engagement between government testing agencies and researchers to explore these questions further.

Public Sector Recruitment in Tamil Nadu

This report focuses on Group Recruitments conducted by the Tamil Nadu Public Service Commission (TNPSC).

TNPSC is one of four recruitment agencies in the state of Tamil Nadu that are responsible for recruiting state-level government employees.[1] Among these state-level recruitment agencies, TNPSC has a privileged status—both because of its long history, and because it is established by the Constitution of India and empowered to conduct its affairs independently of the state government.[2]

The Group Recruitments are by far the most popular recruitments that TNPSC conducts. Between FY 2013 and FY 2017, these recruitments accounted for 94% of all applications received and 81% of all vacancies notified. The Group Recruitments are conducted to staff the rank-and-file administrative posts in the state's various government departments. These exams are popular as they have few eligibility requirements

[1] The other agencies are the TN Uniformed Service Recruitment Board (TNUSRB), which recruits firefighters, police, and other uniformed officers; the Teacher Recruitment Board (TRB), which recruits teachers and other staff for the Department of School Education; and the TN Medical Services Recruitment Board (TNMRB), which recruits staff for hospitals and clinics.

[2] TNPSC has the distinction of being the oldest of the state PSCs, established as the then Madras Public Service Commission in 1929.

and are known to recruit for particularly desirable posts.

TNPSC follows a recruitment procedure that is similar to many of the other state PSCs in the country. Each recruitment goes through the following process:

The Recruitment Process

1. **Notification**. For candidates, the process starts when the Commission issues a notification. A notification is a legally binding document that provides the full terms and conditions for the recruitment.
2. **Registration**. Once the notification is released, candidates have 1-2 months to apply for the exam. Candidates apply using an online form.
3. **Examination**. Depending on the type of exam, recruitment may proceed in either a single stage, or in three stages. For single-stage recruitments, candidates sit for a single multiple choice test. For the multi-stage recruitments, candidates are filtered successively through three stages: first a multiple choice preliminary exam; then an essay-based main exam; and finally, an oral interview with a panel of judges. In multi-stage recruitments, selection is determined based on the sum of the main exam score and interview marks; the preliminary exam does not count towards the final score. Oral test marks can take only one of five values, and, by the Supreme Court's ruling in *Mohinder Sain Garg v. State of Punjab [1990]*, are capped at 15% of the total score.
4. **Selection**. TNPSC can recruit for several thousand vacancies in a single Group Recruitment. Once the scores are finalized, the top-scoring candidates select the vacancy they most prefer, in rank order and according to the reservation rules.

   To match candidates to vacancies, TNPSC conducts what is known as its "counseling" process. In counseling, the top-scoring candidates—about 2-3 times the number of available vacancies—come to the TNPSC office in Chennai to participate. (Those who do not come forfeit their post.) During counseling, candidates are called in rank order and asked to choose among the remaining vacancies that apply to their reservation category, as shown on a large screen.[3] The whole process is recorded on video and in conducted in public.

[3] Candidates can also choose not to take a post. This may happen, for example, if the candidate already has a government job and prefers to keep it over choosing one of the remaining options.

There are five different types of Group Recruitments that TNPSC conducts. These differences are summarized in Table 0.1 below.

Table 0.1: Types of Group Recruitments Conducted through TNPSC

	Group 1	Group 2	Group 2A	Group 4	VAO
# Exams in Sample	4	3	3	5	2
Minimum Age	21	21	21	18	21
Minimum Qualification	UG	UG	UG	10th Std	10th Std
Selection Rounds	3	3	1	1	1
Vacancy Range	74 - 139	1,064 - 1,241	1,807 - 2,269	4,963 - 9,351	813 - 2,342
Select common posts	Deputy Collector, DSP	Revenue Assistant, Inspector	Assistant	Junior Assistant, Typist, Stenotypist	Village Administrative Officer

Notes: The table covers Group recruitments notified between FY 2013 and FY 2019. The combined Group 4/VAO exams conducted in FY 2017 and FY 2019 are included in the Group 4 category. The combined Group 2/2A exam is included in the Group 2 category.

Abbreviations: UG = Undergraduate Degree; DSP = Deputy Superintendent of Police

Group 1 recruits for the highest level of the bureaucracy within the state cadre.[4] Group 4 and VAOs are the most junior officers recruited through state-level exams.[5]

An Overview of the Data Used in this Report

There are three main sources of data that are used throughout the report.

Administrative Data

TNPSC maintains a set of databases that track candidates as they go through the recruitment process. These datasets will be collectively referred to as Administrative Data. This data

[4] Group 1 officers report to officers selected through the Union Public Service Commission (UPSC) exam. After an extended period of service, Group 1 officers can be promoted to UPSC-level Service, such as the Indian Administrative Service.

[5] The Village Administrative Officer (VAO) is responsible for maintaining village land records. In other states, these officials are known as *patwaris* or *talathis*, among other names.

is available for all Group Recruitments conducted between FY 2013 and FY 2019. In total there were 17 Group Recruitments conducted during this period.

The Administrative Data includes:

- **Application data**: This is information that candidates submit to TNPSC when they fill out the online application.
- **Exam performance**: The data provides the total number of marks candidates have obtained at each stage of the exam process.[6] In some cases, I also observe which candidates answered which question correctly.
- **Selection**: This data indicates which candidates selected a post through counseling.

[6] Preliminary score marks were not available for the 2016 Group 1 recruitment, but we can identify which candidates appeared for the exam.

Candidate Survey

Although the administrative data covers the universe of applicants, it provides limited information on candidates' experience with the exam process. To complement the administrative data, I ran a survey of 3,574 candidates who were preparing for a future TNPSC exam. The survey data provides rich information on candidates' social and economic background, their beliefs about the exam process, and their investments in exam preparation—all variables that are not found in the administrative data.

The survey was conducted between May 2022 and July 2022—after the notification for the 2022 Group 4 exam was released. Consequently, 90% of our sample consists of candidates who were planning to appear in that year's Group 4 exam, which was conducted on July 24th.

The survey was conducted entirely online, and respondents filled out the survey on their own. Respondents were recruited mostly through Facebook advertising. To encourage participation and survey completion, candidates were promised a Rs. 50 Amazon gift card and/or study material.

At the end of the survey, candidates made a choice of which gift they wanted to receive: either 1) the Rs. 50 Amazon gift card; or 2) a soft copy of an information card shown to candidates as part of the survey. In addition, *all* candidates, regardless of their gift choice, received a link to an online practice test that we created, which replicated actual past TNPSC Group exams.[7]

[7] To prevent fraud, we conducted a verification procedure candidates after the survey was complete. See Appendix A.2 for details.

Follow Up Phone Survey Between October 2022 and January 2023, I conducted a follow-up survey with 1,749 candidates who said they were planning on participating in the Group 4 exam. In contrast to the baseline survey, this one was conducted over the phone by an enumerator.

The follow-up survey was designed to capture both how candidates' economic activity may have changed after they took the exam, and what their expectations might be for future exams.

Sampling weights Survey respondents are not sampled in a representative manner. As a result, the sample looks different from the population as a whole.

Table 0.2 compares the characteristics of survey respondents with the broader population. Given that the data came from a convenience sample, the Candidate Survey does a remarkably good job of matching the variation in geography, caste, gender, age, and marital status. The major difference between the two groups is in their educational profile: survey respondents are more likely to be college graduates, and less likely to have completed their schooling under the Tamil Nadu State Board.

To adjust for the sampling bias, I re-weight observations according to their relative frequency of certain candidate characteristics in the population.[8] All estimates in the report that use the Candidate Survey data use these weights, unless otherwise noted.

[8] I re-weight using a set of variables whose distribution has proven to be stable over time, namely age, community, and college attainment.

Interviews and Focus Groups

At times, I quote candidates directly. These quotes are based on a set of 3 interviews/focus groups that were conducted with 5 TNPSC candidates in February 2022. These conversations were hosted online and recorded with participants' consent.

In other parts of the report, I relay summaries of conversations. These are references to the various interactions I have had with candidates appear for state-level public service commission exams over the years—both in Tamil Nadu, and in other states. These interactions have shaped my thinking in important ways, and wherever possible I try to convey the lessons I have learned in this report.

Table 0.2: Comparing the Administrative Data and the Candidate Survey

Variable	Admin Data Average	Survey Data Average
Panel A: Demographics		
Age	26.3	24.1
Female	0.552	0.543
Unmarried	0.674	0.703
Panel B: Education		
College graduate	0.573	0.720
10th Std Board: TN State Board	0.892	0.791
Panel C: Reservation Category		
Unreserved Community	0.020	0.030
SC/ST	0.271	0.205
Destitute Widow Quota	0.004	0.024
Ex-Serviceman Quota	0.003	0.016
Disability Quota	0.013	0.029
Panel D: Geographic Variation		
Chennai District	0.040	0.070
Northern Districts	0.206	0.172
Eastern Districts	0.229	0.237
Western Districts	0.243	0.231
Southern Districts	0.275	0.288

Notes: The Admin Data Average is based on candidates who had a valid hall ticket in either the 2019 Group 4 exam or the 2019 Group 1 exam. This column is a weighted average of the share of candidates applying in Group 4 and the share applying exclusively in Group 1, in the same proportion that these two groups of candidates appear in the survey. The Survey Data Average is unweighted.

District Grouping:

- *Northern Districts = Dharmapuri, Krishnagiri, Ranipet, Chengalpet, Kancheepuram, Tirupathur, Vellore, Tiruvannamalai, Thiruvallur*
- *Eastern Districts = The Nilgiris, Coimbatore, Tiruppur, Erode, Dindigul, Karur, Namakkal, Salem*
- *Western Districts = Viluppuram, Kallakurichi, Cuddalore, Perambalur, Pudukottai, Ariyalur, Thanjavur, Nagapattinam, Mayiladuthurai, Thiruvarur, Trichirappalli*
- *Southern Districts = Madurai, Kanyakumari, Thirunelveli, Tenkasi, Tuticorin, Ramanathapuram, Virudhunagar, Sivagangai, Theni*

Part I

Understanding Candidate Application Behavior

A TNPSC candidate's stack of study materials

1 Who applies?

1.1 Motivation

TNPSC conducts the most popular state-level public sector recruitment exams in Tamil Nadu

There are more candidates who appear for the TNPSC Group 4 exam than there are candidates appearing for any other state-level public sector recruitment exam. Table 1.1 tabulates the application volume for some of the most popular recruitments in the other state-level agencies; TNPSC's application volume is nearly four times as large.

Table 1.1: Application Volume in TNPSC vs Other Select Public Sector Recruitment Exams, FY 2019

Exam	Recruitment Agency	Appeared Candidates
Teacher Eligibility Test (Paper II)	Teacher Recruitment Board	3,79,735
Combined Grade II Recruitment	Uniformed Services Recruitment Board	$\approx$ 3,25,000
Nurses Recruitment	Medical Services Recruitment Board	57,789
Group 4	TN Public Service Commission	13,66,548

Data Sources: TRB: TRB Website (Accessed Jan. 2023). TNUSRB: Vijay Kumar (2020). TN MRB: TN MRB Website (Accessed Jan. 2023).

To get a sense of scale, TNPSC receives about as many applications in Group 4 just from Tamil Nadu as the Central Governemnt's Staff Selection Commission receives from *the entire country* to fill posts with similar job descriptions.[1]

TNPSC application volume can also dwarf the size of online job recruitment platforms in the state

For example, consider the following comparison between TNPSC and a popular low-wage online job recruitment platform. This comparison is instructive because there is likely meaningful overlap between its user base and TNPSC's applicant pool (see Table 1.2).

[1] This refers to the SSC's Combined Higher Secondary Level Exam, in which 14,09,856 candidates appeared for the Tier I exam in 2019 (SSC Annual Report 2020-2021).

Table 1.2: Comparing Average Characteristics of Jobseekers in TNPSC and in an Online Platform

Variable	TNPSC	Platform
Female	0.55	0.35
Age 18-24	0.53	0.48
College grad.	0.57	0.69

In a typical week in 2019, the platform attracted about 42,000 jobseekers and hosted about 2,500 full-time vacancies in Tamil Nadu.[2] This means there were roughly 16.8 jobseekers for each full-time vacancy in Tamil Nadu. If the platform had as many vacancies as TNPSC notified in FY 2019, then we would expect it to see 1.11 lakh jobseekers on the platform. *By comparison, 14.5 lakh unique candidates appeared for a TNPSC Group Recruitment in FY 2019.*

[2] The data for the online platform is available at: "G2LM|LIC - How Labor Market Tightness and Job Search Activity Changed in the First Year of COVID-19 in India: Evidence from a Job Portal." Research Data Center of IZA (IDSC). Version 1.0. doi:10.15185/glmlic.707.1.

Does TNPSC's large footprint translate into market power?

Organizations have market power when their actions can coordinate the behavior of a sizable share of the market.

Whether TNPSC has market power depends in part on how applications are distributed across the population. The larger the share of a given labor market that participates in TNPSC recruitments, the more likely that private sector firms in that market will need to contend with TNPSC policy in their own hiring decisions.

Labor markets are typically segmented by age, education, geography, and gender. Consequently, these are the key axes along which we will study heterogeneity in participation in TNPSC exams.

What does the size and composition of the applicant pool tell us about the state of the labor market?

Every recruitment is usually accompanied by a headline about the extraordinary number of highly qualified individuals appearing for posts with minimal qualifying requirements. This is often taken to be a sign of distress in the labor market. The implicit assumption is that if private firms had robust labor demand, then jobseekers would be less likely to appear for public sector recruitment exams.

But whether or not this is assumption is true is an open question. Another possibility is that the high application volume reflects ambition rather than desperation. Acquiring a job that provides lifetime security and prestige against all odds is no small feat. As the economy grows, perhaps it is also the case that more and more people have the wherewithal to aspire, rather than feel compelled to take the first job available?

Before we can draw inferences about the state of the labor market from changes in TNPSC application volume, we first need to understand what those changes *mean*.

1.2 Application Trends

This section puts TNPSC's application volume in historical context by presenting key facts about how the size and composition of the applicant pool has changed between FY 2013 and FY 2019.

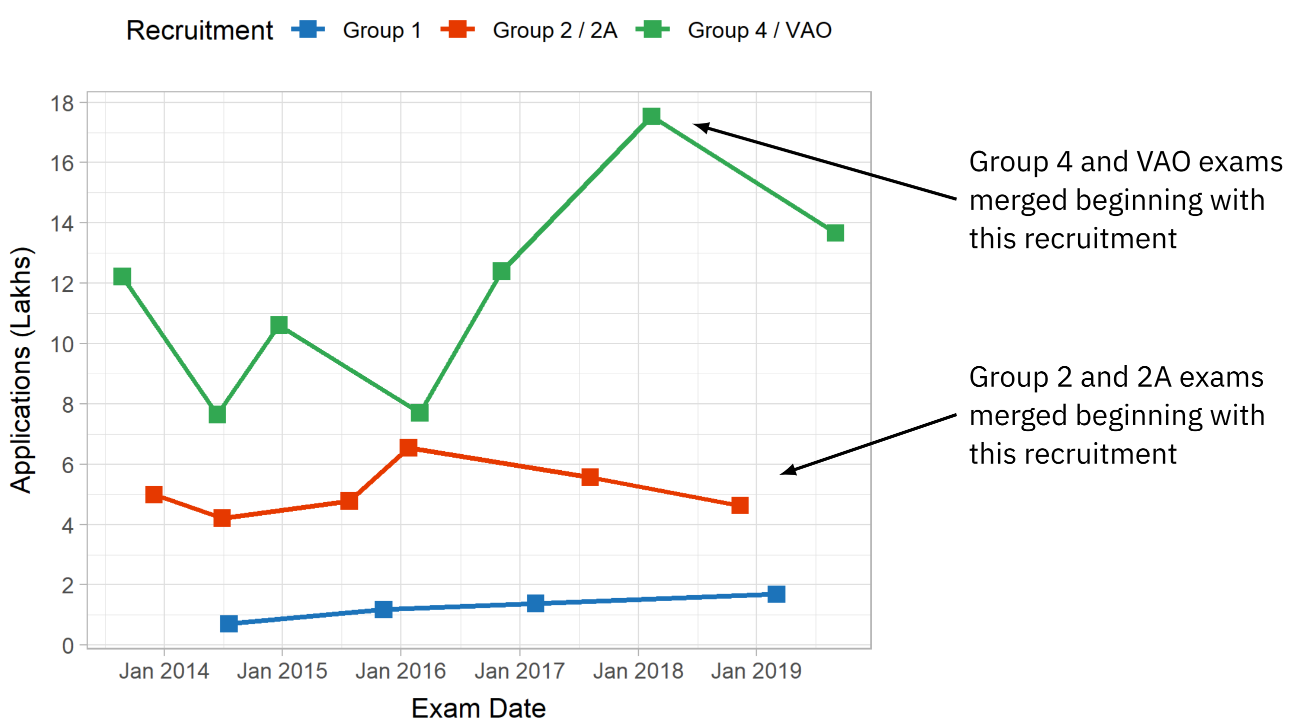
Changes in Application Volume

At each level of recruitment, the number of candidates appearing has either increased over time or remained steady

In Group 1, there has been a steady increase in the number of candidates appearing for the exam. A total of 98,002 more candidates appeared in FY 2019 compared to FY 2013, which, once one factors in the time that elapsed between these two exam dates, corresponds to an average increase of 21,204 candidates per year.

In Group 2/2A, the number of candidates appearing has remained roughly constant throughout the sample period.

Figure 1.1: Trends in application volume by Group



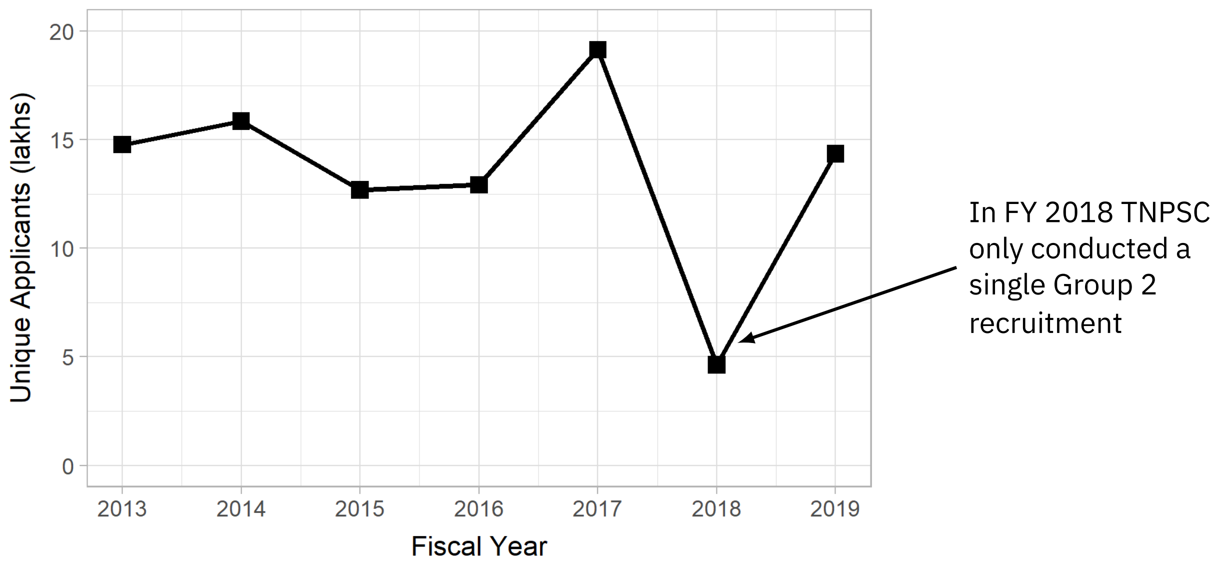
Notes: Candidates can apply to multiple exams, so these categories are not mutually exclusive.

In Group 4/VAO Recruitment there was a large increase in application volume starting in FY 2017. Part of this is explained by the merger of the Group 4 and VAO exams, but not entirely. The number of candidates appearing in FY 2017 and FY 2019 is still 94–120% higher than the combined number of unique individuals appearing for the then-separate Group 4 and VAO exams in FY 2014.

However, there has been no change in the number of candidates appearing for TNPSC Group Recruitments overall

This is because of changes in cross-exam application patterns. Once one takes into account the duplicate candidates appearing across exams within the same fiscal year, we find that there are about 15 lakh unique individuals appearing for TNPSC Group Recruitments in almost every fiscal year (see Figure 1.2).

Figure 1.2: Trends in Total Unique Applicants



Given that educational attainment and population size have been increasing over this time period, the size of the eligible population has almost surely expanded as well.[3] This means that application rates in the population are almost surely falling.

[3] Precise figures are not available because the last Census was conducted in 2011-2012.

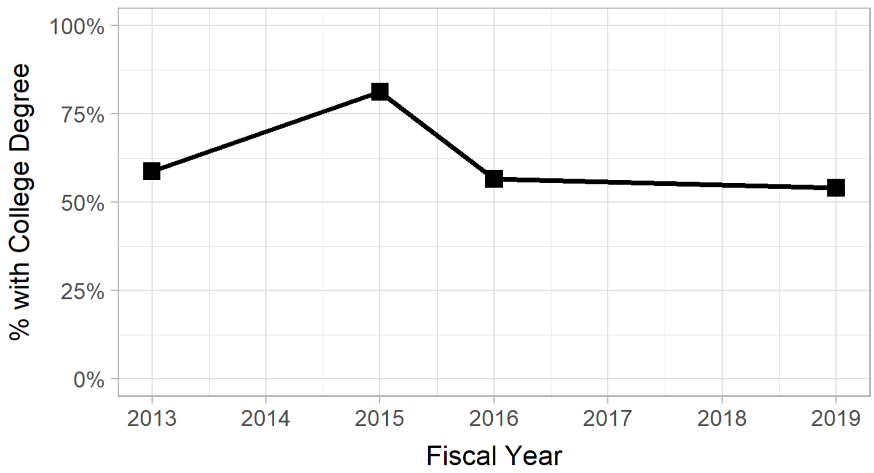
Changes in the Composition of the Applicant Pool

Although the total number of candidates has stayed constant, the composition of candidates has been changing significantly. Below I highlight some noteworthy changes.

There has been a marked increase in the share of candidates with engineering degrees

The applicant pool is about evenly split between those with a college degree and those without. This ratio has mostly remained stable between FY 2013 and FY 2019 (see Figure 1.3).

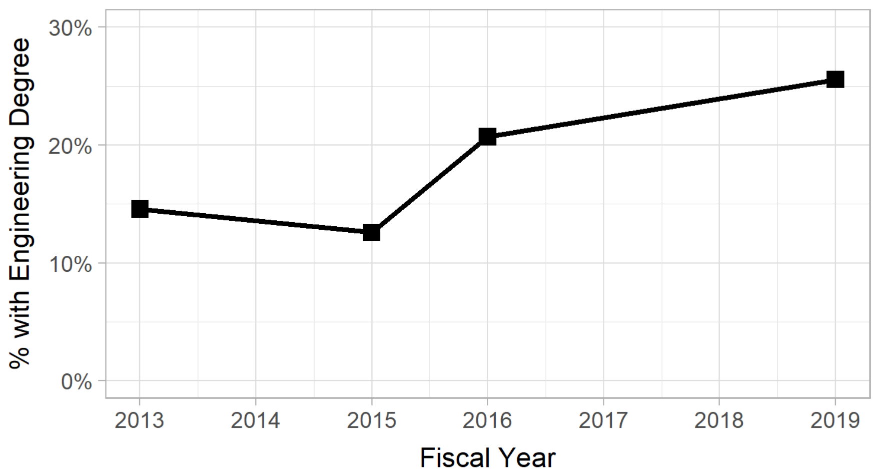
Figure 1.3: Share of TNPSC candidates with a College Degree



Notes: For comparability, this figure restricts attention to fiscal years in which a Group 1 notification was notified

However, there has been a change in the *types of degrees* that we see among TNPSC applicants.

Figure 1.4: Share of College Graduates with Engineering Degree in the TNPSC Applicant Pool



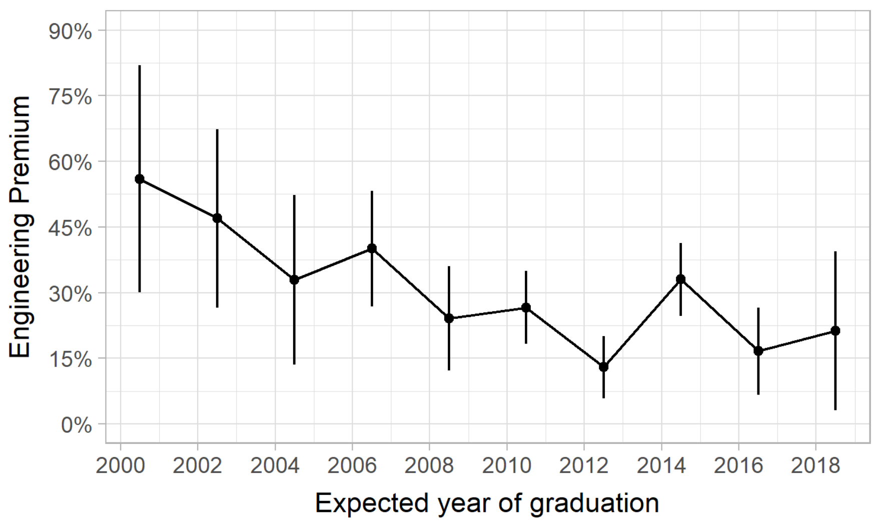
Notes: For comparability, this figure restricts attention to fiscal years in which a Group 1 notification was notified

In particular, the share of college graduate applicants with either a Bachelor of Engineering (B.E.) or Bachelor of Technology (B.Tech.) degree has nearly doubled over this time period (see Figure 1.4).[4]

This trend is noteworthy because engineering degrees have traditionally had relatively strong private sector job prospects compared to other degrees.

However, this is becoming less true with time. Figure 1.5 plots estimates of the wage premium accruing to engineering graduates in Tamil Nadu by cohort. Over the course of the past two decades, the engineering wage premium has fallen by half.

Figure 1.5: The declining engineering wage premium in Tamil Nadu



Data Source: CMIE Consumer Pyramids Household Survey
Notes: Estimates of the wage premium are based on Mincer regressions that include a quartic in experience and fixed effects for caste and religion. The sample is restricted to men with an undergraduate degree (but not higher). Cohorts are combined in groups of two for more precise estimates. The bars plot 95% confidence intervals.

The decline in the engineering wage premium may be linked with engineering graduates' increasing participation in TNPSC exams. In my interactions with them, candidates who had an engineering degree often talked about their disappointment with both their campus placement, and with private sector job opportunities more generally. This suggests that public sector recruitment may have emerged as a new way for engineering graduates to try to maintain the wage premium that they had initially expected. Whether this is the case in general is a question that merits further investigation.

[4] The trend has been noticeable enough that the coaching centers instructors I interacted with often commented on it, based on their observation of the kinds of students enrolling in their classes.

There has been dramatic progress in women's representation in the applicant pool over time

Figure 1.6 plots the change over time in the ratio of female to male applicants by district. This figure is based on the number of unique individuals appearing for *any* TNPSC Group Recruitment in that fiscal year.

The gender ratio is classified as follows:

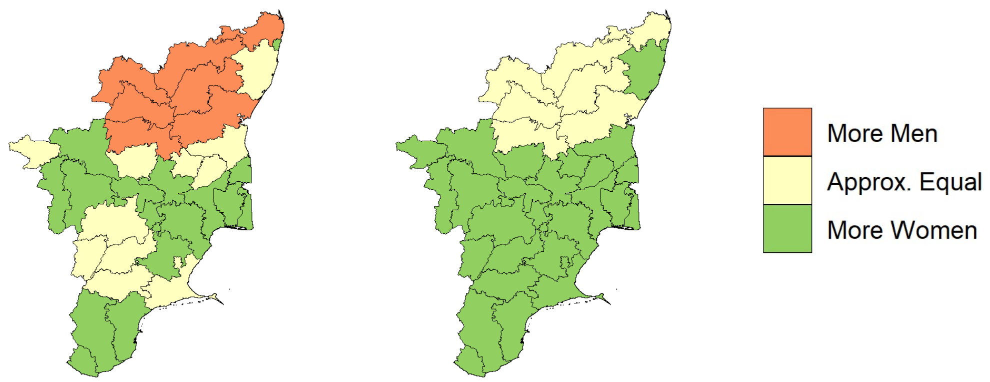
- Districts in which the female-to-male ratio lies between 0.9 to 1.1 are marked as having "Approximately Equal" representation.
- In case the female-to-male ratio is below 0.9, then the district is marked as having "More Men"
- If the female-to-male ratio is greater then 1.1, it is marked as having "More Women."

The graph on the left plots this ratio for exams notified in FY 2013. The graph on the right plots the same for exams notified in FY 2019.

Figure 1.6: Application gender ratio by district, FY 2013-FY 2019

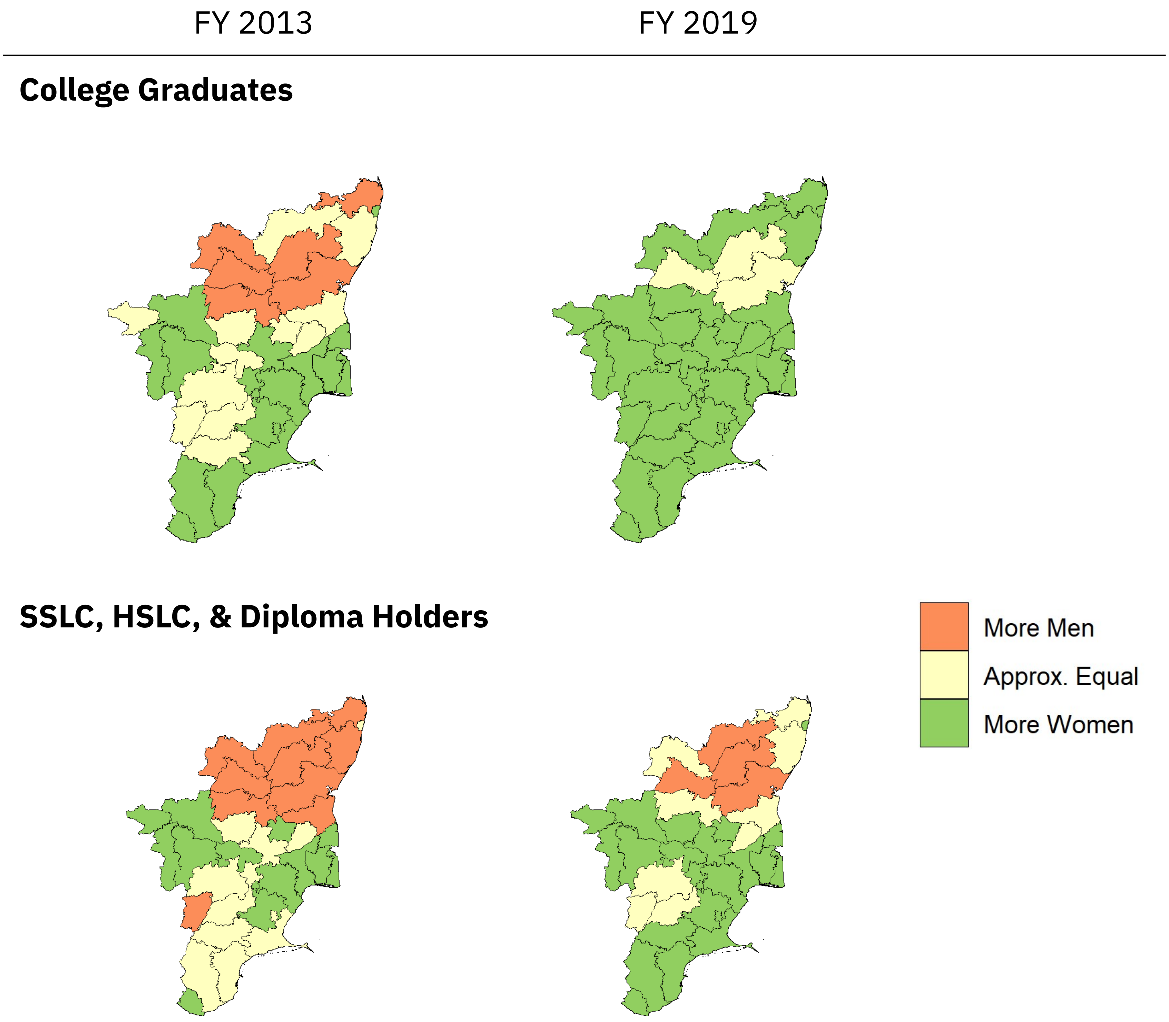
FY 2013	FY 2019

Overall



In FY 2013, there was substantial variation in women's representation in the applicant pool. By FY 2019, women started to outnumber men in almost all districts, and where they did not their representation was approximately equal.

Figure 1.6 (cont.): Application gender ratio by district, FY 2013-FY 2019



Among non-college graduates, there are still some districts where men meaningfully outnumber women. But only in the Northern districts of Dharmapuri, Thiruvannamalai, Vellore, and Villupuram.

Note that women's representation in TNPSC Group Recruitments has been increasing even though women's participation in the workforce overall has not. According to the latest government surveys, the female labor force participation rate in Tamil Nadu in both rural and urban areas has remained essentially stagnant over this time period, if not falling (Pandey, 2021). The fact that women continue to participate in TNPSC recruitment suggests that women may be willing to work if provided the right opportunities.

1.3 Where are applications concentrated?

This section studies how application rates vary along four axes: age, education, geography, and gender.

To compute application rates, I compare application counts in the TNPSC administrative data with population totals based on the 2011 Census. By focusing on the recruitments conducted closest in time to when the Census data was collected—namely the FY 2013 recruitments—I minimize the measurement error that results from extrapolating population projections far into the future.[5] The downside, of course, is that these estimates of application rates are a little dated.

[5] I account for the 2 year gap between the TNPSC Admin data and the Census data by extrapolating the 2001-2011 average growth rate forward by two years.

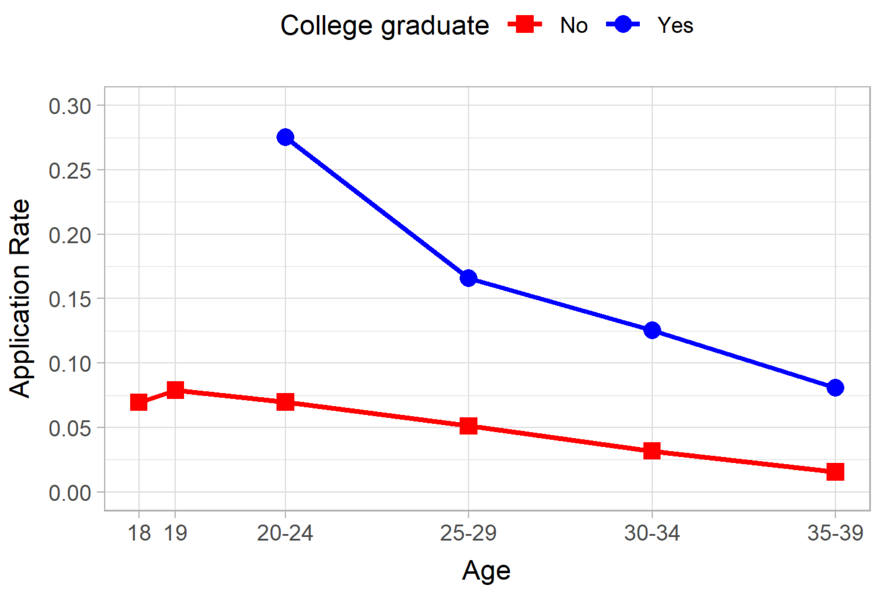
Application rates peak for recent college graduates, and decline with age.

Application rates for recent college graduates are extremely high. In FY 2013, *about 1 in every 4* fresh college graduates in the entire state of Tamil Nadu appeared for a TNPSC Group recruitment.

The application rate among college graduates is much higher than that of non-graduates—-so much so that even college graduates in their late 30s are still participating at rates much higher than any non-graduate.[6].

[6] Still, about half of the applicant pool does not have a college degree (see Figure 1.3). This is because there are so many more people without a college degree to begin with that a lower percentage of them applying still results in roughly equal representation.

Figure 1.7: Application rates in FY 2013 by Age and Education

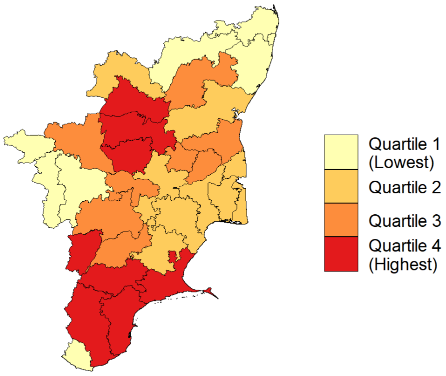


Application rates vary substantially across the state

In FY 2013, the district with the highest application rate was Dharmapuri, where *approximately 22% of all eligible individuals* between the ages of 18 to 35 participated in a TNPSC Group Recruitment.[7] Virudhnagar, Theni, and Thirunelvi all had participation rates between 15 to 20%. On the other end of the spectrum, the districts of Chennai, Thiruvallur, and The Nilgiris had participation rates of about 4-5%.

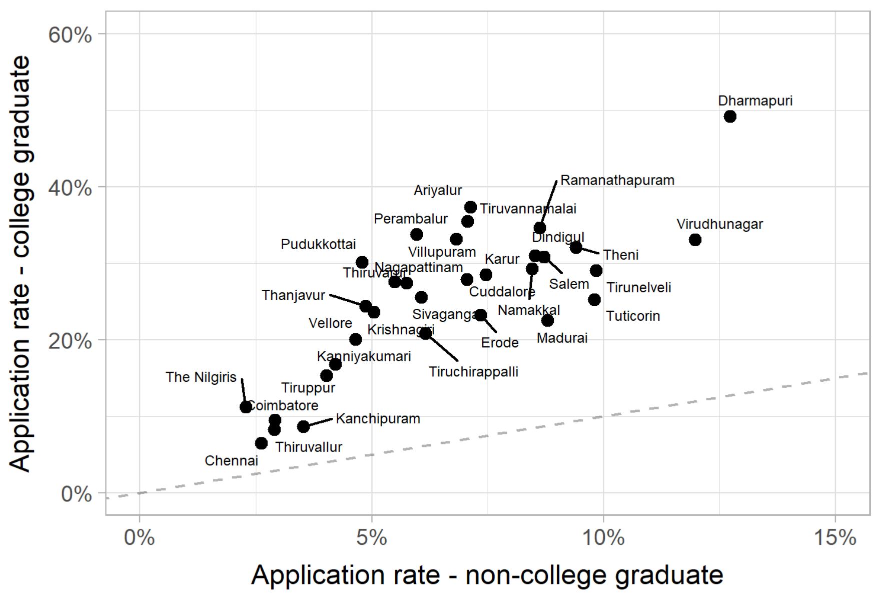
[7] Candidates are assigned to districts based on their reported permanent address.

Figure 1.8: Application rates in FY 2013 by district among 18-35 year old eligible individuals



The districts with the highest application rates among college graduates also tend to have the highest application rates for non-college graduates.

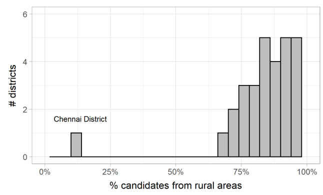
Figure 1.9: Application rates within districts are correlated across education levels



Notes: The dashed line marks the 45 degree line.

Within districts, most applicants are from rural areas (see Figure 1.10). Candidates are identified as coming from a rural area if their permanent address is located in a pincode that has Branch Office of the Indian postal system. By this definition, Dharmapuri and Pudukkottai districts have the highest share of candidates from rural areas. Aside from Chennai, Sivaganga district has the lowest share.

Figure 1.10: Share of candidates within districts from rural areas



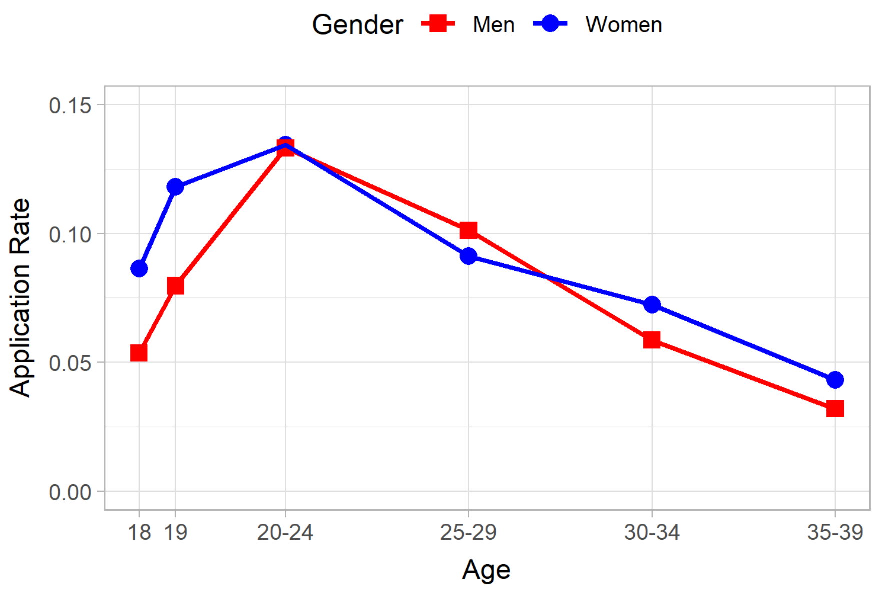
Women apply at higher rates then men, except during the main marriageable and child-bearing years

In Figure 1.11, we see that across most age groups, women participate in TNPSC Group Recruitments at par with or at a higher rate then men.[8]

[8] Recall, this data is from FY 2013, when men still outnumbered women in a large part of the state (see Figure 1.6).

Figure 1.11: Application rates in FY 2013 by Gender

(a) Overall



However, women's participation falls below men's between the ages of 25 to 29. Why might that be the case? One factor that seems to be important is the differential pressure women face to get married and act as caregivers to their families.

In the Candidate Survey, we asked candidates why they think they might stop preparing for the exam. About 16% of unmarried women below the age of 25 said they were likely to stop because of marriage pressure. By contrast, only 4% of comparable men answered the same way.

Why are women so keen to apply?

In focus group discussions with female TNPSC candidates, three main themes emerged:

1. It gives them a sense of agency

   This sentiment is perhaps not unique to public sector employment, and has common features with the reasons why women want to participate in the labor force more generally.

   > *"Q: Suppose you got selected in TNPSC. How do you think your life will change?*
   >
   > *A: ...I think it will change completely. Because when you are a housewife...mmm....there is nothing. But if we are in a government job, there will be financial support. And I see that my friends who have government jobs have the confidence to do all their work by themselves."*

2. They expected their future in-laws to let them continue working only if they had a government job

   As one candidate put it:

   > *"If I continue working in a private job after my marriage, my in-laws won't put much priority on it. But if I get a govt job, they can't force me to leave or quit my job; hopefully no one will say that."*

   Given these expectations, women who want to have a career after marriage may feel like their only chance of doing so is securing a government job.

   Some women go to great lengths to resist the pressure to abandon their career goals. One candidate relayed the following story about her friend:

   > *"One friend told me that in order to avoid a marriage proposal, she completely shaved her head, without telling anyone in advance. Thus the marriage was postponed, and now she is preparing for exams."*

3. They prefer the timing and work culture in the public sector

   Government jobs are amenable not just to in-laws, but to candidates themselves. They often talked about how the timings and lack of work pressure in the public sector made it possible for them to balance their responsibilities at home and at work.

Figure 1.11 (cont.): Application rates in FY 2013 by Gender

(b) College Graduates

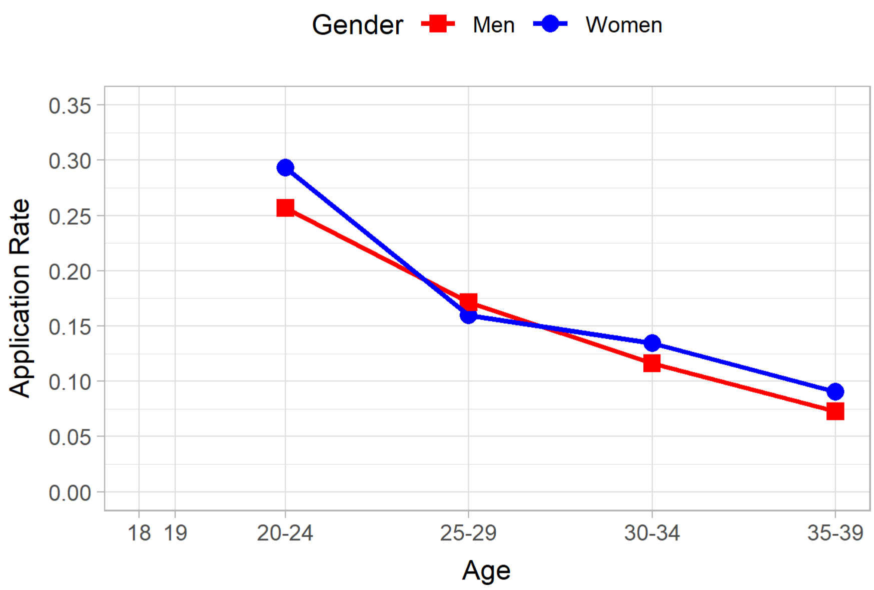
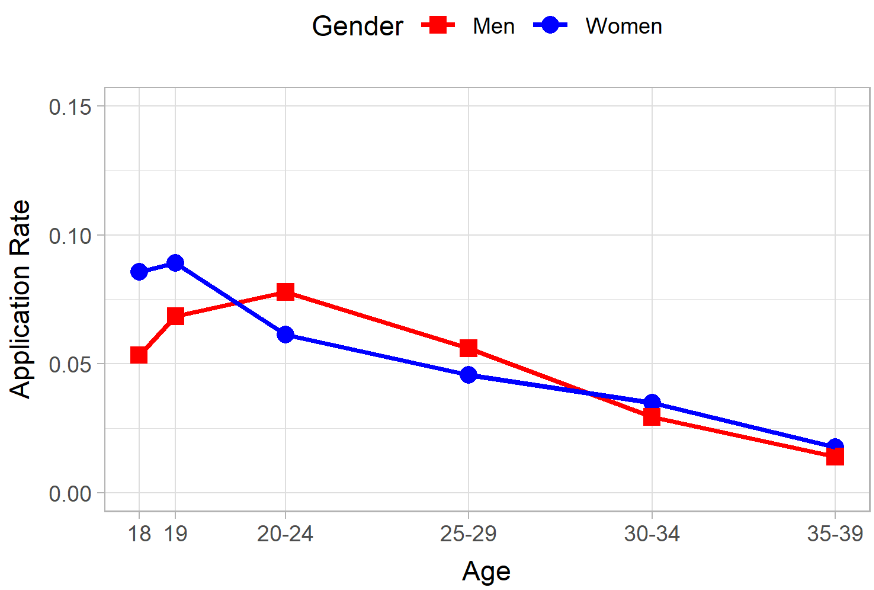


Figure 1.11 (cont.): Application rates in FY 2013 by Gender

(c) SSLC, HSLC, and Diploma Holders

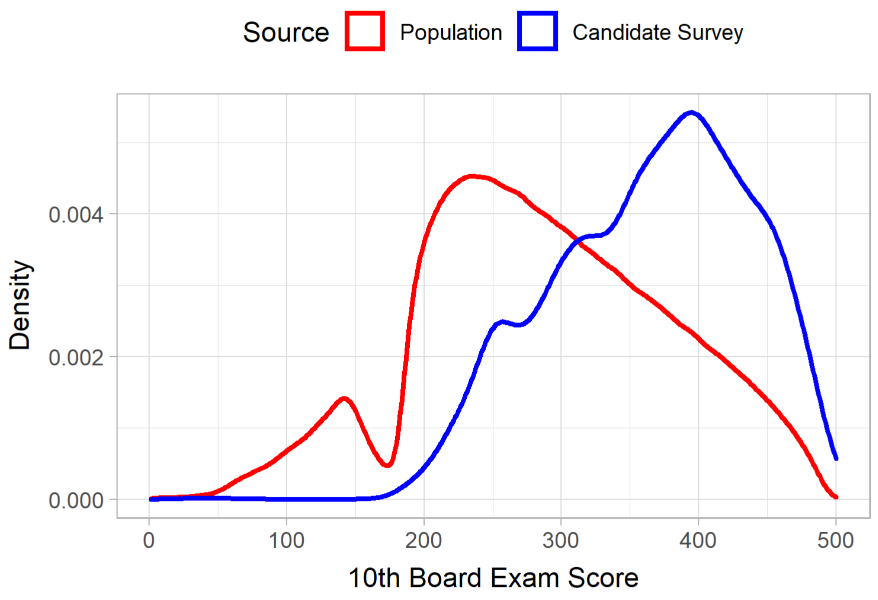


Higher-performing students are over-represented in the applicant pool

In the FY 2019 Group Recruitments, 90% of candidates completed their schooling under the Tamil Nadu State Board.

How do candidates' board exam scores compare with the population as a whole? In the Candidate Survey, we asked respondents to report their 10th exam board marks. We then compare this distribution to the distribution of marks obtained by all students who participated in a recent 10th standard board exam.

Figure 1.12: TNPSC Candidates tend to have done well in school



Notes: The figure restricts attention to individuals who participated in the Tamil Nadu State Board exam

Top-scoring students are over-represented in the applicant pool. For example, there are about 3.5 times as many candidates scoring 90% and above in the TNPSC applicant pool compared to the population overall (13% versus 4%).

About 80% of the unemployed population in the state applies through TNPSC

In the 2019 Group 4 Recruitment, TNPSC asked applicants to report their employment status as part of their application for the first time. According to this data, there were 1,257,650 unemployed candidates who appeared for the exam.

Meanwhile, according to the Centre for Monitoring the Indian Economy, there were about 1,560,000 unemployed individuals in Tamil Nadu at the time the application was live.[9]

[9]NB: this estimate depends on extrapolating 2011 Census numbers forward about a decade.

Putting these statistics together, I obtain a rough estimate that approximately 80% of all unemployed people in Tamil Nadu are preparing for TNPSC exams.

Does reducing vacancies affect unemployment rates?

If so many people are applying for jobs via public sector recruitment, does changing the availability of public sector jobs affect the unemployment rate?

Between 2001 and 2006, there was a natural experiment on this question in Tamil Nadu. In these years, due to a state financial crisis, TNPSC implemented a partial hiring freeze that resulted in a 86% drop in vacancies.

As it turns out, during this period, applications *increased* by 7%, and employment rates among male college graduates *dropped* by 13%. The effect can't be driven by the fact that fewer people got government jobs, because the drop in employment is about 20 times as large as the reduction in the number of vacancies. This means that the reduction in vacancies affected candidates' willingness to work, at least within this sub-population.

In theory, candidates could have responded to the hiring freeze either way. They could choose to give up on full-time exam preparation and take up employment in the private sector, since they didn't know when the next exam would be. But they could also choose to double down on exam preparation, both to compete for the remaining vacancies and to be ready when vacancy levels returned to normal. In that case, candidates could end up spending more time out of work.

The general takeaway is that it is in fact possible to affect unemployment rates by changing recruitment policy. But forecasting what those effects will be is challenging because it likely depends on candidates' expectations on the availability of future vacancies.

For a more detailed analysis of the hiring freeze, see Mangal (2022b). "The Long-Run Costs of Highly Competitive Exams for Government Jobs."

1.4 Desperation or Aspiration?

A key question is whether candidates are applying because they are unemployed, or whether they are unemployed because they are applying. In the former case candidates are struggling to find any work, and applying to TNPSC as part of that struggle. In the latter case, candidates have an alternative employment option available, but choose not to take it because they aspire to acquire a job within the state government services.

In this section we aim to better understand candidates' intentions for applying. We start by studying how application behavior responds when candidates are exposed to better economic opportunities.

Areas with more economic development have lower application rates.

It is commonly believed that participation in government recruitment exams is a sign of under-development, because it shows that there are no other decent jobs available. Is this in fact the case?

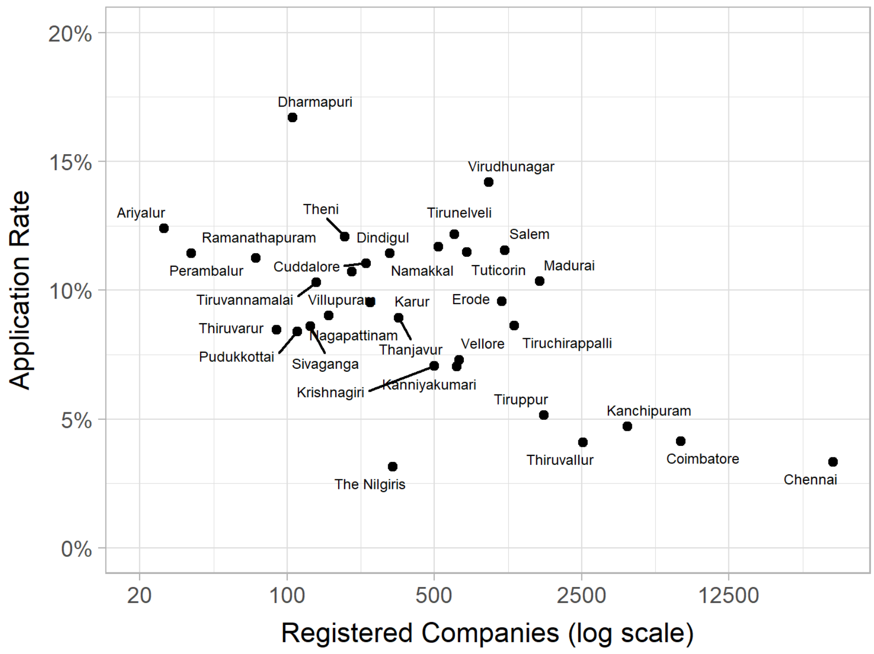
To answer this question, I look at how application volume varies with the presence of large firms across districts in Tamil Nadu. Large formal firms are in the best position to offer employees good salaries and benefits.[10] One might wonder whether candidates who apply to TNPSC are those who either lack exposure or access to these private sector alternatives.

[10] For a more detailed discussion of the wage premium associated with large firms and multinationals, see: Colonnelli et al. (2018), Hjort et al. (2020), and Arellano-Bover (forthcoming).

I measure the presence of large firms using the number of companies registered with the Ministry of Corporate Affairs with addresses in that district as of March 2015. Candidates are assigned to districts based on their permanent address. Our outcome is the share of eligible individuals between the ages of 18 and 35 who wrote any TNPSC Group exam in FY 2015.

Figure 1.13 shows that there is a clear downward slope. The industrialized zones of Chennai and Coimbatore have among the lowest application rates. Meanwhile, districts that are known to lack industry, such as Ariyalur and Ramnathapuram, have some of the highest application rates.

Figure 1.13: Application Rates and the Presence of Registered Companies



However, when there are temporary improvements in local economic conditions, applications increase

Recall that most candidates are from rural areas (Figure 1.10). In rural areas, where most families still depend directly or indirectly on agriculture for their livelihood, periods of good rainfall create additional job opportunities. By comparing districts that received an unexpectedly good period of rainfall (relative to their own past history) with the other districts in the same time period, we can better understand the link between economic activity and application behavior.[11]

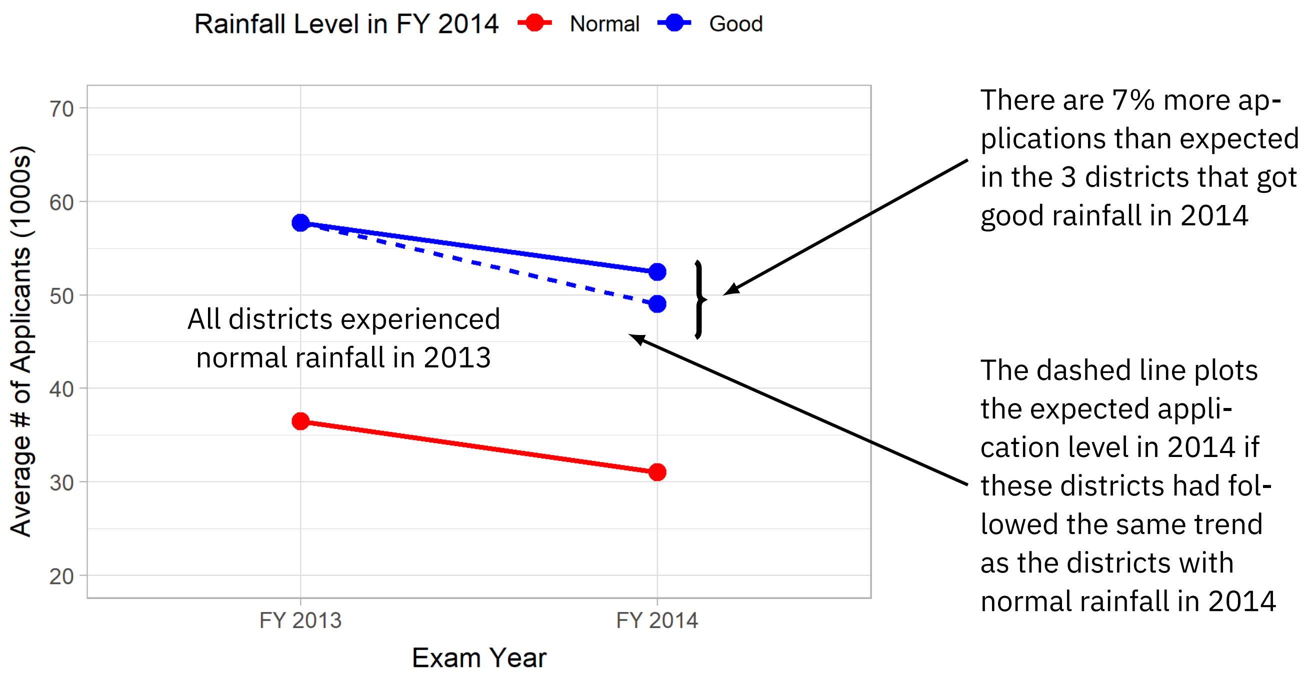
[11] In the analysis below, "good" rainfall is defined as rainfall in the 80th percentile of the observed rainfall distribution in that district between 1960 and 2011.

Good rainfall can cause application volume to either go up or down. Since there are better job opportunities available, candidates may be more willing to start working. On the other hand, local economic booms also provide households with more income. That extra income may be used to help finance additional exam preparation, leading to more applications.

A key advantage of using variation *within* Tamil Nadu is that we can control for other potential confounding factors. There are many reasons why the application volume will vary from recruitment to recruitment, including the number of vacancies offered, the timing relative to other recruitments, and national macroeconomic conditions. We can control for all of these confounding factors by comparing application levels among districts *within the same recruitment*, which is conducted on the same day for everyone.

Figure 1.14 compares two sets of districts. All the districts experienced normal rainfall in the year preceding the 2013 Group 4 exam. However, in the year preceding the 2014 Group 4 exam, there were 3 districts that experienced above-normal rainfall (specifically, Thoothukkudi, Tirunelveli, and Virudunagar). Under the assumption that which districts got good rainfall is essentially random, we can measure the causal effect of a small improvement in local economic conditions on application behavior.

Figure 1.14: The effect of good rainfall on application volume, Group 4



It turns out that periods of good rainfall are associated with *more applications for government jobs*, not less. And if we expand our sample to include more recruitments and rainfall shocks we end up with the same conclusion.[12]

This tells us that, at least in the short run, rural economic growth, with its accompanying improvement in local labor market opportunities, goes hand in hand with TNPSC exam participation.

[12] The full details of this analysis will be available in a forthcoming academic working paper.

Candidates as Constrained Dreamers

One way of making sense of these patterns is to look more deeply at the candidate's perspective. What emerges is a portrait of candidates as *constrained dreamers*, where both desperation and aspiration co-exist. Most candidates have very difficult economic circumstances, and are short on funds. But the dream of a government job is deep-rooted, and candidates do not settle for their alternatives easily. Exam participation is therefore often more a question of whether it is feasible and practical to continue to pursue the dream rather than a question of willingness.

The aspiration to obtain a government job is near universal among Tamil Nadu government school students

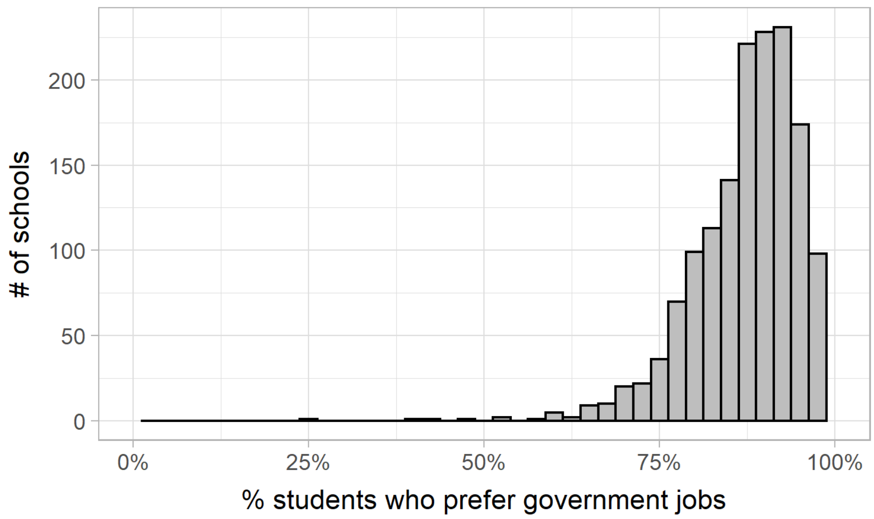
In 2022, the Directorate of School Education conducted a survey of 12th standard students to better understand their career aspirations. The survey included the following question:

> *If you had to choose between working in one of the following, which one would you prefer the most?*
> 1. *Working in a Government Job*
> 2. *Working in a Private Job*
> 3. *Running my own Business*

On average, 88% of students said they would prefer a government job.[13] Moreover, the vast majority of students *within* each school said the same as well:

[13] In the same survey, students were also asked whether they intend to write any competitive exam in the next four years, including TNPSC Group Recruitments. About 66% of respondents said they would.

Figure 1.15: Fraction of students in Tamil Nadu Government Schools who prefer a government job



A corollary of this graph is that there is very little variation in preference for government jobs at the district level—despite the fact that, as we saw in Figure 1.13, there is substantial variation in the private sector opportunities across districts. This means that the presence of large private firms in one's district does not bend students' ambition away from the public sector.

Instead, the students who are among the least likely to prefer a government job are those who do not plan to pursue higher education, with only 65% preferring government jobs compared to 88% of the remaining students.

But most candidates have few resources available to help them achieve their goals

In the Candidate Survey, a majority of respondents reported a household income of less than Rs. 10k per month. For reference, as of April 2022, Rs. 9,770 to Rs. 10,658 per month was the official state minimum wage.

Table 1.3: Candidates' household income

Household Income (per month)	Share (%)
Less than Rs. 5,000	18.8
Between Rs. 5,000 - 10,000	36.1
Between Rs. 10,000 - 20,000	27.4
Between Rs. 20,000 - 30,000	10.2
More than Rs. 30,000	7.5

One could imagine how, given the level of competition, only those with sufficient access to resources would choose to participate. But that turns out not to be the case.

Most candidates anticipate that if they drop out it will be because of lack of money, not a lack of motivation

In the poorest households, about four times as many candidates say that they would stop for lack of money than for lack of motivation.

Table 1.4: Candidates' household income and their anticipated reasons for stopping exam preparation

Household Income (per month)	Lack of Money (%)	Lack of Motivation (%)
Less than Rs. 5,000	36.9	9.2
Between Rs. 5,000 - 10,000	41.4	10.4
Between Rs. 10,000 - 20,000	39.1	13.5
Between Rs. 20,000 - 30,000	37.6	21.1
More than Rs. 30,000	28.6	47.8

Notes: Respondents were allowed to specify multiple reasons.

Although there is financial need, candidates are not willing to take up just any job either

In the Candidate Survey, we asked respondents the following question:

> *If you were not studying for competitive exams, what job would you be doing now?*

It is striking that about one quarter of candidates say they would not be working if they stopped preparing.

Table 1.5: Candidates' anticipated outside employment option

Alternative Occupation	Women (%)	Men (%)	Overall (%)
Private Employment	40.4	52.5	46.5
Daily wage labor	8.8	16.1	12.5
Farming	6.7	12.3	9.5
Business	6.9	7.4	7.2
Unpaid work / Not working	37.1	11.7	24.3

Notes: Columns sum to 100%.

Of course, planning not to work immediately after dropping out of exam preparation will likely play out very differently for men and women. It's reasonable to expect that most of these men will join the labor force eventually, since labor force participation among educated men is still nearly universal. For women, given the low rates of employment in later years, there is a good chance that women who drop out of exam preparation will never join the labor force.

But in both cases, the fact that there is hesitation to join the labor force outside of selection through TNPSC suggests that candidates exercise discretion when choosing which jobs they will take.

The candidates who hesitate the most to start working after exam preparation also likely have the most job opportunities available

Here we focus on men, for whom family permission to participation in the workforce is not a confounding factor.

College graduates are slightly more than twice as likely to say they would not work compared to non-graduates. Furthermore, it is not the case that non-graduates anticipate disguising their unemployment by withdrawing into self-employment. Instead, they see themselves participating in private sector jobs at substantially higher rates than college graduates.

Table 1.6: Male candidates' anticipated outside employment option by college attainment

Alternative Occupation	College (%)	Less than College (%)
Private Employment	43.4	60.9
Daily wage labor	22.5	10.2
Farming	11.8	12.6
Business	6.0	8.7
Unpaid work / Not working	16.3	7.5

Notes: Columns sum to 100%.

These patterns emerge despite the fact that we would normally think of college graduates as having more employment opportunities available than non-graduates. (After all, this is the reason why people generally go to college in the first place).

The most likely explanation is that college graduates are more particular about finding a job that matches their level of education—a sentiment that we not only heard in our own interactions with candidates, but which also has been well-documented in ethnographic studies of candidates preparing for government job exams (Jeffrey, 2010).

Summing up

Candidates are generally poor and unemployed. But that doesn't mean that they are willing to take any job they can find. A sizable share of the candidate pool would not work at all if they were not preparing. And when private sector opportunities expand, we see an increase in applications, not fewer. What seems to matter is not just whether *any* job exists, but whether the jobs that are available are good, such as the kind that tend to be provided by large, formal firms.

This could be a good thing. It can mean that candidates have enough support that they can afford to dream and no longer have to take the first job they find. But it also suggests that reducing the queues will be challenging. It also means that, in the meantime, candidates' aspirations will become more and more frustrated.

1.5 Discussion

So far we have seen that there are many sub-populations within Tamil Nadu where TNPSC application rates are very high. In this section, we consider the ways in which these high application rates can result in recruitment policy having broad influence over the rest of the economy.

Public sector recruitment policy could affect skill development in the state

The most prestigious government jobs require a college degree. Given how common it is to aspire for a government job, it is possible that there are some people who decide to go to college just because it helps meet this eligibility requirement.[14] Similarly, the choice of major or degree type may also depend on students' expectations about applying to TNPSC in the future.

[14] In Taiwan, Xu and Adhvaryu (2022) find this to be true. Due to the nature of their reservation policy, some people had a higher baseline probability of success in the civil service exam. Then in 1980 the government reformed the policy in a way that ended this advantage. The authors find that the people who lost their advantage in the exam process were less likely to complete both high school and college after the reform.

This raises the possibility that public sector eligibility requirements can be used as a policy lever for encouraging college completion. But whether this is desirable to do, even based on the limited criteria of economic efficiency, is an open question.

The answer depends on the kinds of skills or knowledge that students on the margin of enrolling would gain in college. Previous studies have found that students on the margin of attending engineering colleges and elite colleges typically benefit from doing so (Bertrand et al., 2010; Bagde et al., 2016; Sekhri, 2020). It is possible the effect generalizes to less prestigious institutions as well.

On the other hand, colleges and candidates could end up responding to this policy in counter-productive ways. For example, in Bihar it is common for candidates to enroll in a college in the state in name only while they live in Patna full time to attend coaching classes. Even aside from such extremes, a perverse equilibrium could develop where students are happy not learning in college because it allows them more time to focus on competitive exams; and colleges are happy with low attendance, because it is less work for teachers and the administration.

Another potential consequence of the high application rates is that the syllabus that is set for TNPSC exams can have far-reaching consequences on the skill base that is available to private sector firms. Most candidates are never selected. But many of them will have spent some time preparing for the exam in some way. It is possible that the types of skills that the current exam rewards are the types of skills that will become more common in the labor market in the near future.

If the government wishes to encourage candidates to develop a particular skill—e.g. computer word processing—then including that skill on the TNPSC exam may jump-start the process. However, adding “difficult” skills to the exam also runs the risk of thwarting the government’s other goal of making sure the pool of selected candidates reflects the diversity of the state, including marginalized populations.

One way of potentially resolving this issue is to lean more heavily on reservation to fulfill the representation goals. This might then free up the Commission to use the syllabus for other objectives, including adding questions that encourage candidates to build less familiar but useful skills.

Public sector recruitment policy could affect wages and employment

We have seen that in some districts young college graduates participate in TNPSC Group Recruitments at rates upwards of 30% (see Figure 1.9). With participation rates this high, it is not hard to imagine how TNPSC policy decisions could induce shifts in labor supply that could have important consequences for wage levels and employment rates. After all, there is precedent. The rates at which young college graduates participate in TNPSC Group Recruitments is similar to the rates at which rural households have participated in the Mahatma Gandhi National Rural Employment Guarantee Scheme (NREGA). Furthermore, there is a rich literature documenting how high participation rates in NREGA can have major spillover effects on the village economy and beyond (Narayanan and Das, 2014; Imbert and Papp, 2015; Muralidharan et al., 2016, 2017).

How exactly recruitment policy will affect labor market outcomes is unclear. It depends on how candidates react and firms adjust, both questions where more research is needed before any definitive statements can be made.

Should the government aim to reduce application volume?

At present, most young people are chasing dreams that few will have the chance of realizing. This is likely to be a cause of deep psychological distress for many people.[15] We do not yet have data that tell us the full extent of this problem. But it is troubling that there are regular appearances of headlines about people committing suicide when they are unsuccessful in PSC recruitment exams.

[15] Jeffrey (2010) provides an account of how candidates who prepared for the selection exams and were not successful developed deep psychological wounds, including long-lasting inferiority complexes and feelings of being left behind and worthless that likely affected their ability to reach their full potential even after they stopped studying.

Nonetheless, reducing application volume for its own sake is likely to be counter-productive. The government should be careful to avoid an approach that tells students they should simply give up on their dreams. In addition to generating an explosive reaction, this approach is also a waste of young people's energy. It is a blessing to have a society where people are ambitious and willing to take risks to achieve their goals. The problem is not that students are ambitious; the problem is that they are all ambitious for the same thing.

What could we do instead? The patterns in candidate application behavior suggest some possible solutions.

The value of career counseling and exposure Recall, application rates are highest in areas where students have likely not seen other aspirational careers besides government service. This suggests that it may be useful to invest in programs that help students develop new and more diverse sets of aspirations, and, crucially, to provide them with a clear pathway for achieving these alternative goals.

A potentially powerful way of introducing students to new career possibilities is through media, which has been shown to have a powerful influence on how people make major life decisions.[16] The Tamil Nadu government has already recognized some of the challenges that students face in navigating careers outside of government, and has launched a campaign called *Naan Mudhalvan* to help. It may be worth exploring ways to use the scheme's platform to target students who intend to apply for public sector recruitments but are also unaware and uncertain of other careers.

[16] See, for example: Ferrara et al. (2012) discuss how soap operas shape fertility decisions; and Riley (2019) shows how a movie about a woman African chess champion encouraged girls to perform better in exams. La Ferrara (2016) provides an overview of this literature.

Reforming the private sector The preference for public sector jobs may be as much a question of avoiding the private sector as it is about the seeking the advantages of the public sector.

A key issue that policy needs to address is that searching for jobs in the private sector is currently fraught with danger. When we asked TNPSC candidates what we as a research team could provide that might help them look for jobs in the private sector, a common response was some form of help in distinguishing which ads are genuine.

The problem is that there are some ads designed to defraud job seekers. Very organized scammers will put up a fake ad, conduct a bogus interview, offer a fake job, and then require some kind of upfront payment to cover "training" or some other made-up expense. After jobseekers make these payments, the "firm" vanishes.[17]

[17] For more details on how these fake job scams happen, see Poonam (2021).

By cracking down on on fraudulent job posting, the government may be able to encourage more young people to participate in the private sector labor market. In turn, by decreasing search costs for genuine firms, this kind of reform may also improve the quality and quantity of jobs offered.

2 How much do candidates invest?

2.1 Motivation

What are the total resources of time, money, and effort that get absorbed in the recruitment process?

A government job recruitment does not begin when vacancies are notified. It begins a year or more in advance when candidates decide to start studying (in some cases full-time) in order to maximize their chances of success whenever the next exam is announced.

There is no other entry-level non-specialized white-collar job in India where this kind of extended investment in job search is expected or normal. Of course, public sector recruitment has its own exigencies. But even taking those into account, is it really necessary for so much time, money, and effort to be spent in selecting the most meritorious candidates for government jobs? Or are there a way of organizing the recruitment process that produces matches that are just as good (if not better) at a lower cost to society? What are the potential economic gains from improving the efficiency of the current system?

To answer these kinds of questions, we must first understand what the costs of the current system are. The *social cost* of a recruitment is the sum of all the resources that people put into the exam process. This includes not just the cost of organizing and conducting the exam, but also the costs that people bear in the process of competing.

Candidates bear a wide range of costs in the process of competing, including: attending specialized coaching classes; buying special books and other study materials; renting space in private study halls;[18] migrating to parts of the state that have turned into hubs for civil service exam preparation;[19] and forgoing income in order to dedicate more time towards exam preparation.

In this chapter, our goal is to understand: i) how large the costs of recruitment are; ii) how they are distributed between candidates and the government; iii) and how the candidates' share of the cost is distributed in the population.

[18] For candidates who prefer not to study from their room or home, they can rent out space in a private study hall, which offers candidates a quiet space where they can concentrate on their studies.

[19] Across Tamil Nadu there are well-known hubs for TNPSC exam preparation that attract candidates from surrounding areas. Some of the biggest ones include Anna Nagar, Chennai, and Kancheepuram.

2.2 The social cost of recruitment

This section presents estimates of the *total* social cost of recruitment, using the 2022 Group 4 recruitment as a case study.

In this analysis, the social cost of recruitment has three main components:[20]

1. **Exam Costs**: There is a cost associated with organizing and conducting the exam itself.
2. **Direct expenditure by candidates**: This includes the expenses that candidates make for the sole purpose of helping them improve their performance on the exam.
3. **Opportunity Cost**: By spending time on exam preparation, candidates give up time that they could otherwise spend working. The opportunity cost of exam preparation is the income that candidates forgo in order to focus on exam preparation. To make time to study, candidates may work less than they otherwise would, or they may take up less demanding (and hence less remunerative) jobs than they otherwise would.

I infer Exam Costs from TNPSC's budget. I estimate both the Direct expenditure and the Opportunity Cost using data from the Candidate Survey. Throughout this chapter, all analyses using the Survey data are restricted to the sample of respondents who were planning to write the 2022 Group 4 exam.

There were approximately 18.5 lakh candidates that appeared for the 2022 Group 4 exam.[21] I calculate total costs by multiplying my estimates of the average cost per candidate by 18.5 lakh. Since the Candidate Survey is reasonably representative of the overall population (see Table 0.2), this approach should provide us with estimates of the right order of magnitude.

I then compare the social cost of recruitment against the value of the jobs themselves. This gives us a measure of the efficiency of the recruitment process. That is, we want to know, for every Rs. 1,000 the government spends on salaries for new jobs, how much does society as a whole spend on figuring out how to allocate people to those jobs?[22]

Obviously, there are benefits to screening as well, so just looking at the costs is not enough to make a conclusion on whether the current system is inefficient. However, as we will see, the costs of recruitment are currently very high. It is worth look-

[20] This accounting is not complete. For example, I do not account for the resources spent by either government departments or private organizations to support TNPSC candidates. These expenditures are much harder to track because they are so decentralized—but I expect that these expenditures are much smaller than the ones I do capture.

[21] TNPSC has yet to release an exact figure for the number of candidates appeared. This figure is sourced from an article in The Hindu (published July 24, 2022).

[22] This is similar to the concept of "deadweight loss" or "excess burden" of taxation. The cost of raising tax revenue must also take into account the fact that people change their behavior in response to the tax. Similarly, the cost of offering jobs with above-market wages is not just the wages themselves, but also the potential losses stemming from the resources that people expend competing for those high wages.

ing into whether there are ways of reducing these costs without sacrificing other recruitment objectives.

The value of the jobs offered in the recruitment

The total lifetime value of salaries offered in the jobs listed in the 2022 Group 4 exam is approximately Rs. 3,664 crore.

I arrive at this number by calculating the Net Present Value (NPV) of the expected salary progression based on the pay band listed in the notification.[23] The NPV calculation takes into account the full stream of future payments resulting from obtaining a post, and includes basic pay and the House Rent Allowance.[24] For the NPV calculation, I fix the interest rate at 6.75%, which is the current 1 year fixed deposit rate at the State Bank of India. The pay increments associated with each year of service are provided in the current pay matrix (see TN Finance Dept. G.O. Ms.No.90, Dated 26 Feb. 2021).

I use data from the previous Group 4 recruitment to predict the length of service for selected officers. The average candidate was 28 years old when they were selected in the 2019 Group 4 recruitment. Given the current mandatory retirement age of 60, this implies that selected candidates typically undertake 32 years of service. I therefore calculate the value of the posts offered in this recruitment over the course of a 32 year career, assuming that candidates stay within their pay band and get the prescribed annual increment.[25]

[23] A copy of the notification is available on TNPSC's website at `https://tnpsc.gov.in/English/Notification.aspx`. The notification number is 07/2022.

[24] The House Rent Allowance HRA) depends on the city in which the employee is posted and the basic pay level. To simplify the calculation, I take the HRA to be an extra 5% of base pay, which is roughly the median amount offered at each pay level. See TN Finance Dept. G.O.Ms.No.305, Dated 13th Oct. 2017

[25] I do not include Dearness Allowance, assuming that it will keep track with actual inflation on average. This means that our estimates are measured in *real* (not nominal) 2022 Indian Rupees.

Table 2.1: Expected lifetime salary outlay for jobs notified in the 2022 Group 4 Recruitment

Pay Scale	# Vacancies	NPV per Vacancy (Lakh Rupees)	Total NPV (Crore Rupees)
Grade Pay 1650	49	42.31	20.73
Grade Pay 2000	1	47.29	0.47
Grade Pay 2400	6,220	49.83	3099.40
Grade Pay 2800	1,031	52.69	543.24
Total	7,301		3,663.85

Notes: The value of each vacancy is calculated over a 32 year career. Future income is discounted based on a 6.75% interest rate. Monetary values are measured in 2022 real Indian Rupees.

The cost of conducting exams

The cost of conducting the Group 4 exam is approximately Rs. 91 to Rs. 103 crore.

TNPSC's 2019-2020 Annual Report provides the latest publicly available information on how much it might cost to run a Group 4 exam.

That year, TNPSC conducted 45 different examinations for 12,777 posts. Its total annual budget was Rs. 88 crore, which is approximately 103 crore in current prices.

The cost of running just the Group 4 recruitment is some fraction of the total budget amount, but it's hard to separate the fixed and variable costs without some assumptions. We can deduce a range of plausible estimates by considering different assumptions on how costs scale. On the low end, if we suppose that costs scale linearly per application received, then the cost of the recruitment would be Rs. 91 crore (in current prices). On the high end, we could suppose that the the cost of recruitment is close to the entire TNPSC budget itself. As we will see, this variation is small compared the other components of the social cost of recruitment.[26]

[26] In the final calculation, it is useful to have a single estimate in the middle of these extremes. I suppose that Group 4 costs twice as much per application received as the remaining recruitments, and arrive at an estimate of about Rs. 97.

Not all of the costs of running the exam are borne by the Government. Through registration and exam fees, candidates shouldered about Rs. 29 crore (in current prices) of total costs, looking at recruitments run by TNPSC. This implies that candidates bear about 28% of the cost of running an exam.

Direct expenditure on exam preparation

Total Direct Expenditure The Candidate Survey asked respondents to report their *lifetime* expenditure on TNPSC exam preparation in four categories: 1) coaching; 2) study material; 3) the cost of staying outside the home while preparing for the exam (including the cost of food and accommodation);and 4) the cost of renting space in a study hall.

To calculate average annual expenditure, I divide lifetime expenditure by the total amount by the number of years spent preparing. This is likely a conservative estimate of the annual expenditure, since candidates may have taken breaks since they started preparing.

Table 2.2: Average annual expenditure on exam preparation

Expenditure Category	Average Annual Expense (Rs.)	Plausible Range (Rs.)
Coaching Classes	634	528 - 740
Books / Study Materials	527	472 - 581
Staying Outside Home While Studying	944	623 - 1,264
Private Study Hall	580	253 - 908
Total	2,687	2,083 - 3,291

Notes: The plausible range reflects the 95% confidence interval of the main estimate, which is a measure of uncertainty that takes into account the sample size of the survey and the variability of responses across respondents

Table 2.2 summarizes candidates' expenditure on exam preparation. Given these estimates, the total expenditure on exam preparation across all candidates is approximately:

$$\text{Total Direct Expenditure} = 18.5 \text{ lakh } \times \text{ Rs. } 2{,}687 = \text{ Rs. } 497 \text{ crore}$$

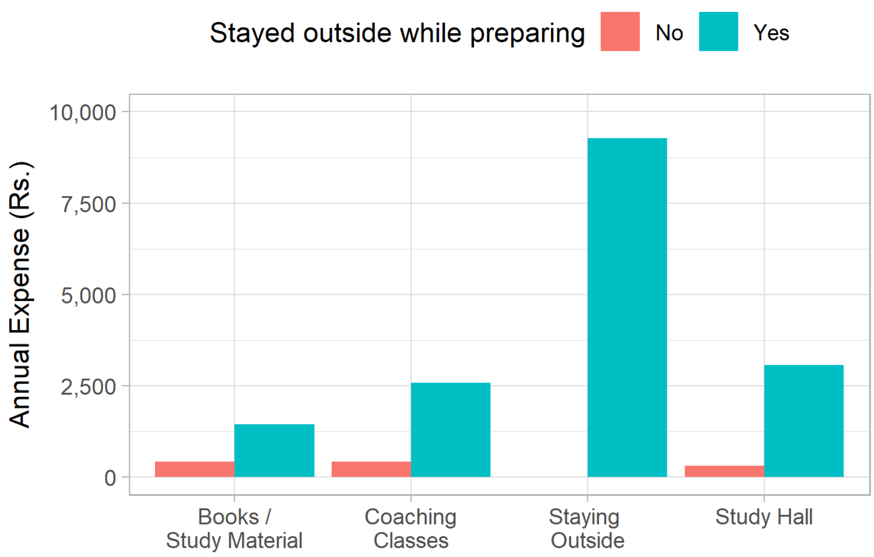
Heterogeneity These averages mask the fact that most expenses are concentrated among a small group of candidates:

- About 44% of candidates do not make any expenses on exam preparation at all (see Table 2.3).
- Even among those who do spend on exam preparation, most of that expenditure is concentrated among the 10% who stay outside of their home while studying (see Figure 2.1).

Table 2.3: Which expenses are the most common?

Expenditure Category	% Making Expense
Coaching Classes	14
Books / Study Materials	51
Staying Outside Home While Studying	10
Private Study Hall	12
Any Expenditure	56

Figure 2.1: Candidates who study outside their home spend more on exam preparation



In total, candidates who stay outside to prepare spend 14 times as much as those who prepare from home. Most of that difference is accounted for by the cost of food and accommodation. But they also spend more in all the other expenditure categories as well.

Since it costs so much to stay outside, why do candidates do it? I asked this question to some candidates in Kancheepuram whose home town is in nearby districts and who do not have any coaching classes to attend during the week. A common response was that home does not provide a good study environment. Candidates tend to get bothered by their relatives at home; many of the men talked of their relatives taunting them "useless" by not earning. Coming to Kancheepuram and studying as a group helps them find support and avoid the negative environment at home.

Why is the market for TNPSC exam preparation relatively small?

We just estimated that the total size of the market for TNPSC exam preparation among Group 4 candidates is about Rs. 500 crore across the whole state. By comparison, experts estimate that coaching centers for civil services *just* in Delhi generate revenue in excess of Rs. 3,000 crore (Vardhan, 2017)—a figure which does not include all the revenue generated by the bookstores, hostels, and other businesses who depend on candidates who come to prepare. This comparison suggests that the direct costs of exam preparation in Tamil Nadu are relatively low.

In our experience, there are three factors that can help explain this difference: 1) the availability of public space to study; 2) the nature of the exam syllabus; and 3) the profile of TNPSC Group 4 candidates.

The availability of public space: In Tamil Nadu there are many public spaces where candidates can study for free, especially in temples and public libraries. In temples, not only is it free to use the space, but they also provide free food at lunch time, which saves candidates money again. Libraries have study materials available, and tend to be filled with candidates around exam time. By contrast, in other states it is relatively uncommon to see candidates use public libraries or temples; they tend to study either at home, in private libraries, or at their coaching academy.

The syllabus: It is widely understood that the exam syllabus in Group 4 draws directly from the syllabus used in school textbooks in government schools (the so-called *Samacheer* books). Thus, many candidates told us that they relied exclusively on school textbooks for exam preparation. Although the full set of textbooks would be expensive to buy, candidates have found ways to avoid purchasing them. For example, candidates loan books from neighbors or friends who no longer need them, or they share books among themselves in study groups. This keeps the costs of study materials down.

Candidates' profile: Group 4 tends to attract candidates from poorer families, without much disposable income. There are, no doubt, some coaching centers for competitive exams that have very expensive facilities—but this is not among centers that specialize in TNPSC preparation. In fact, in Kancheepuram, which is known as a hub for exam preparation in Tamil Nadu, many coaching centers do not have their own classrooms or full-time staff. Instead, they hire already-selected officers who commute from Chennai to give classes on the weekends, and they rent out classrooms in local schools. This suggests that candidates have a small budget, and coaching centers accommodate by finding ways to keep costs down.

The opportunity cost of exam preparation

By spending time on exam preparation, candidates forgo income they could have otherwise earned while working. To calculate this opportunity cost of exam preparation, I use three data points from the Candidate Survey:

- *Current Earnings*: The survey asked respondents whether they were currently working, and if so, how much they usually earn.
- *Alternative Earnings*: The survey then asked respondents what type of job they would do if they were not participating in TNPSC exams. If respondents said they would be working, they were asked how much they expect to earn in that job.[27]
- *Time spent preparing*: The survey asked respondents how many months they have spent preparing for TNPSC Group exams in the past year, and which exams they had taken/were planning to take. This information allows us to infer the amount of time dedicated to Group 4 exam preparation.[28]

Using this information, I estimate the total income that candidates have forgone over the past year while preparing for the Group 4 exam as follows:

$$\begin{aligned}\text{Forgone Earnings} = &\\ &\big(\text{Expected monthly earnings if not studying} - \\ &\qquad \text{Current monthly earnings}\big) \\ &\qquad\qquad \times \text{\# months preparing in past year} \qquad (1)\end{aligned}$$

The main weakness of this approach is that my estimates of Alternative Earnings are based on candidates' expectations instead of observing candidates' actual earnings. If candidates are over-optimistic about their outside option, my estimates of the opportunity cost will be inflated.

As we will see, it turns out that my main conclusions are not very sensitive to how I measure Alternative Earnings. Even with measurement errors as large as 30 or 40% on average, some of the main conclusions will remain robust.

[27] For both current earnings and alternative earnings, respondents reported either weekly or monthly figures, depending on what made sense to them. If respondents reported weekly earnings, the amount was multiplied by 4 to convert it to monthly earnings.

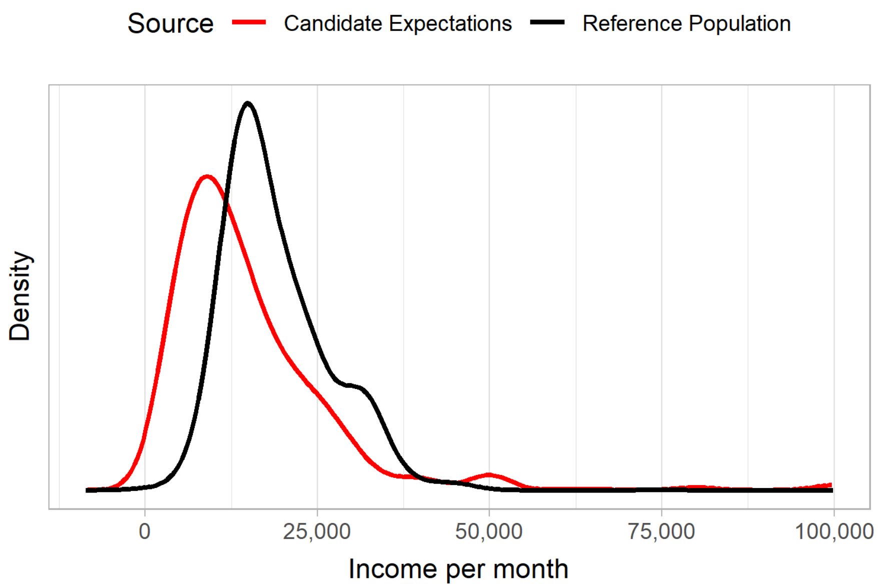
[28] Within the preceding year, the only other Group exam was a Group 2 recruitment. Among candidates who started preparing before the Group 4 notification was released, those who participated in Group 2 prepared on average 0.9 months more than candidates who did not. I therefore subtract this amount from the reported total for all candidates who participated in Group 2 and started exam preparation before the Group 4 notification was released.

Moreover, there are several reasons to think that these figures are not unreasonable:

- For the 25% of candidates who have worked in the past, the median respondent reports the same earnings in their alternative career as they do in their past job.
- On the whole, the responses are within the normal range for what people with similar demographic profiles in Tamil Nadu report in recent labor market surveys (see Figure 2.2).[29]
- In the Follow-up survey, there is a sample of respondents who were not working before the exam but started working after the exam concluded. About 75% of these candidates *under-estimated* their Alternative Earnings, and on average candidates under-estimated their monthly earnings by Rs. 5,301.

[29] The data used in this figure tends to provide relatively higher estimates of earnings than other government surveys (Jha and Basole, 2023). As a result, it is possible that candidates expectations align with market outcome more closely than what is shown.

Figure 2.2: Comparing candidates' expected earnings if they weren't preparing for TNPSC exams with actual earnings in the population



Notes: The figure plots two distributions. The red line shows the distribution that candidates preparing for the Group 4 exam expect to earn if they were not preparing. The black line plots the distribution from a reference population of individuals between the ages of 20 and 30 from the September 2021 wave of the CMIE's Consumer Pyramids Household Survey (CPHS). The CPHS distribution is reweighted to account for the same relative frequencies of job type (i.e. farming, business, daily wage, and other), gender, and college attainment as in the Candidate Survey sample.

Table 2.4 provides estimates of the average value of the main parameters in the opportunity cost calculation.

Table 2.4: The opportunity cost of exam preparation

Parameter	Main Estimate of Sample Average	Plausible Range
Current Income (Rs. / month) [A]	4,968	4,534 - 5,403
Alternative Income (Rs. / month) [B]	11,548	10,822 - 12,274
Expected Monthly Income Loss [A- B]	6,580	5,936 - 7,224
Months spent preparing	4.22	4.06 - 4.37
Forgone Earnings (Rs.)	30,330	26,948 - 33,712

Notes: The plausible range reflects the 95% confidence interval of the main estimate.

The total opportunity cost is therefore:

$$\text{Total Opportunity Cost} = 18.5 \text{ lakh } \times \text{ Rs. } 30,330 = \text{ Rs. } 5,611 \text{ crore}$$

Summing Up

The total cost of recruitment Table 2.5 tabulates the total social cost of conducting the 2022 Group 4 recruitment.

Table 2.5: Social cost of recruitment relative to the cost of vacancies

Line Item	Main Estimate	Lower Estimate	Upper Estimate
Social costs of recruitment			
Cost of conducting the exam	97	91	103
Candidate direct expenditure	497	385	609
Opportunity cost	5,611	4,985	6,237
Total social cost of recruitment (A)	**6,205**	**5,461**	**6,949**
Salary cost of notified vacancies (B)	**3,664**	–	–
Ratio of recruitment to salary costs (A/B)	1.69	1.49	1.9

The estimates tell us that for every Rs. 1,000 that the government spends in salaries, there is an *additional* social cost of Rs. 1,500-1,900 on recruiting candidates for those positions.

Most of this cost does not appear on the government's balance sheet. Instead, it takes the form of reduced activity in the labor market. The part of the overall cost of conducting a recruitment that is visible to the government—namely, TNPSC's budget—is less than 2% of the overall cost of recruitment.

The limits of redistribution through recruitment For some, the sacrifices of exam preparation are worth it as long as public sector recruitment helps redistribute resources and opportunities towards candidates. However, the costs appear to be so large that public sector recruitment turns out to be a very inefficient mechanism for redistribution.

In fact, Table 2.5 tells us that candidates collectively *lose* income by investing in TNPSC recruitment. In other words, compared to how much candidates invest in competing for government jobs, they get less than that in salary from the government in return.

Imagine, hypothetically, that candidates put the money they would otherwise forgo on exam preparation into a common pool, and then allocated it to winners via a lottery (with probabilities that are not necessarily equal). In that case, they would be able to provide an income stream equal to the income stream from a government job for *more* people than what the recruitment offers.

This does not mean that *all* candidates fail to benefit from participating in TNPSC exams. Candidates who are well-informed, well-coached, and who are not overly invested in exam participation may stand to gain on average. But since the total number of vacancies is fixed, and since there seem to be no discernible economic benefits of exam preparation outside of getting a government job, any gain to one candidate will come at the expense of another. This constraint puts severe limitations on the extent to which public sector recruitment can lead to general social upliftment.

Are the costs of competition offset by their long-run benefits?

Whether these costs are worthwhile depends on whether they produce some offsetting social benefits. For example, exam preparation might build skills that help candidates perform better in their job as government employees; it may help them build skills that are useful even if they are not selected; or it may help them obtain a credential of their ability that helps them find a job in the private sector more quickly.

On many of these questions, the answer is just not known. But where there is evidence, it looks like on average long-term exam preparation hurts rather than improves not-selected candidates' earning potential in the private sector labor market.

In general, it is difficult to measure the long-run effects of exam preparation, since candidates who invest more are systematically different from those who drop out early. However, between 2001 and 2006, there was an unexpected hiring freeze in TNPSC that shifted application behavior substantially. As a result, the most exposed cohorts of male college graduates stayed out of the workforce longer, most likely in order to double down on the competition for TNPSC jobs. By comparing the long-run trajectory of these cohorts with the trajectory of the cohorts that preceded them, I can estimate the impact of candidates' investment in remaining on the "exam track" on their long-run labor market outcomes.

In Mangal (2022b), I find no evidence that the candidates who spent more time out of the workforce to prepare for government jobs are better off in the long run. A decade after the hiring freeze ended, male college graduates who were exposed to the hiring freeze have lower employment rates; have lower earning capacities; and have delayed forming their own households. Moreover, the negative effects on candidates appear to spill over to other household members: for example, elder members of the household appear to delay retirement in order to compensate for the income shock.

Together, this evidence suggests that the cost estimates in Table 2.5 may in fact be *understated*, and that the long-run costs of recruitment may be a meaningful component of the overall social cost.

For more details, see Mangal (2022b). "The Long-Run Costs of Highly Competitive Exams for Government Jobs."

2.3 How do candidates spend their time?

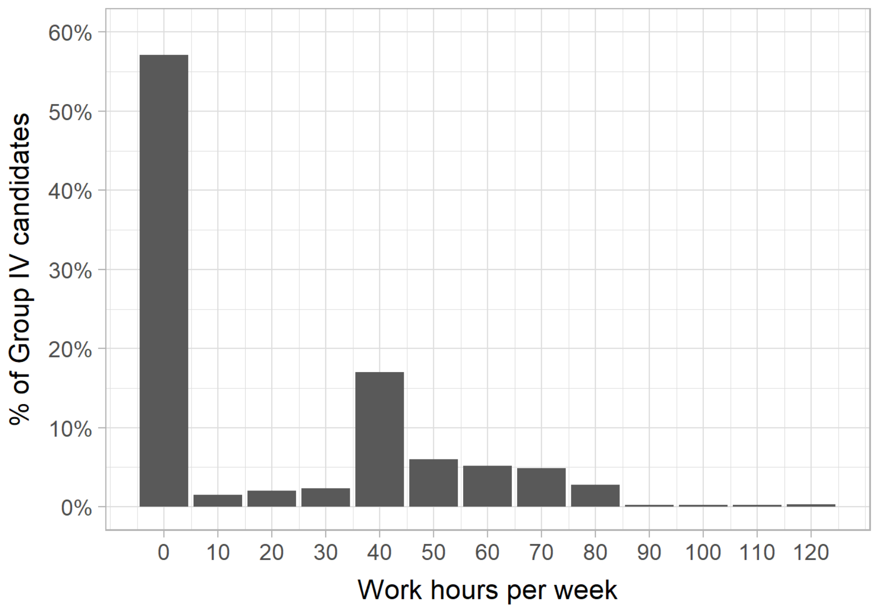
In the previous section we saw that most of the social cost of exam preparation is due to the opportunity cost of candidates' time. Why is this figure so large?

How common is it for candidates to work?

Most candidates who are planning to take the Group 4 exam are not working

Figure 2.3 plots the distribution of the number of hours per week that candidates were working at the time of the survey. Note that the median candidate is not working at all.

Figure 2.3: How much are candidates currently working?



Of those who are working, about 83% are working full-time (i.e. 40+ hours per week). This suggests that either candidates who want to take up part time work have a hard time finding it, or that working part-time isn't worthwhile.[30]

Among candidates who are working, what kinds of jobs do they take? Table 2.6 lists the distribution of jobs that candidates hold by their current work status. Overall, private sector jobs are the most common, followed by daily wage employment. Self-employment is relatively uncommon, except among those working part time.

[30] Part-time jobs or jobs with flexible work hours tend to offer a lower hourly wage rate (see, e.g. Manning and Petrongolo (2008); Goldin and Katz (2011)). This wage penalty may discourage candidates from seeking part-time work.

Table 2.6: Types of jobs held by current Group 4 candidates who are working

	Distribution by Employment Status		
Job Category	Part Time (%)	Full Time (%)	Overall (%)
Private Job	17	63	55
Government Job	4	4	4
Daily wage	43	20	24
Farming	28	12	14
Business	14	11	11

Notes: Each column shows the distribution of jobs held by Group 4 candidates with the given employment status. Respondents could list multiple current jobs, so columns can sum to more than 100%.

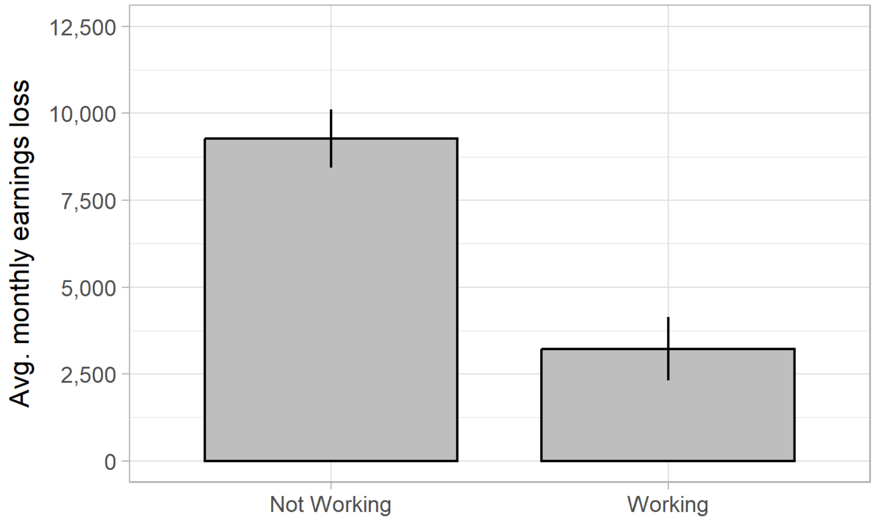
Not surprisingly, candidates who are not working forgo more income per month (see Figure 2.4a). But note that there is still an opportunity cost for candidates who are currently working. This means that candidates who are working while studying expect to earn *even more* if they were no longer planning to take the exam.

Interestingly, among candidates who are working, only men expect to earn more if they stop studying; women on the other hand expect to earn essentially the same (see Figure 2.4b). Figuring out why exactly this is the case deserves future research.

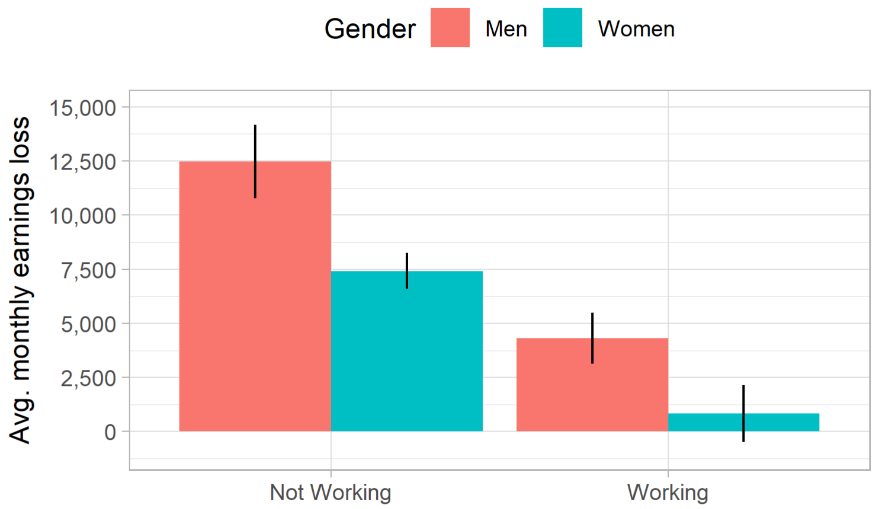
One plausible hypothesis is that candidates choose more flexible jobs and stay closer to home while they are studying. These two priorities reduce candidates' ability to fully capitalize on their skills. Men tend to have more freedom to choose inflexible jobs and look for jobs away from home, so they expect a pay increase when they stop studying. Women are more likely to be confined to flexible jobs that are close to home, so they do not expect to take up better opportunities even when they stop studying.

Figure 2.4: The expected earnings loss due to exam preparation

(a) Overall



(b) By Gender

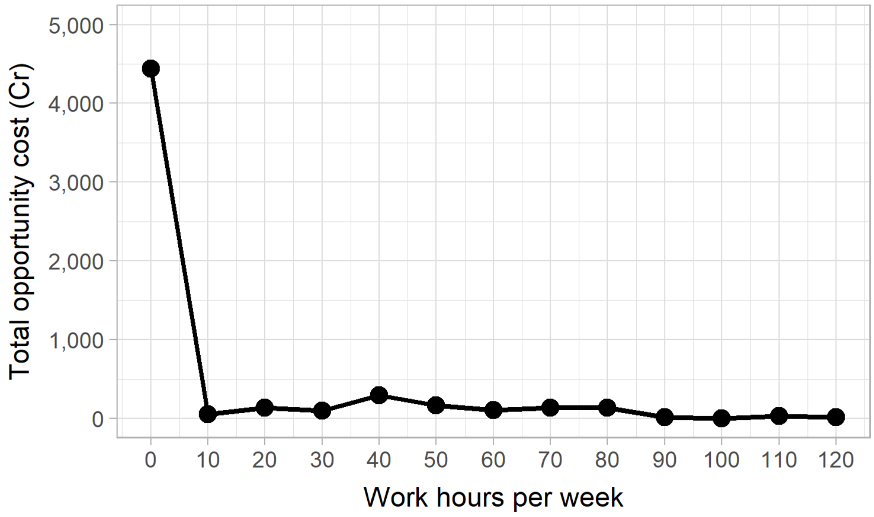


Notes: The black vertical lines indicate the 95% confidence interval for the the estimate.

Most of the total opportunity cost of exam preparation is the result of candidates not working at all

Figure 2.5 divides the total opportunity cost of exam preparation among candidates based on the number of hours they are currently working.

Figure 2.5: The opportunity cost of exam preparation by current work hours



The horizontal axis marks bins of 10 hours. The vertical axis is the total opportunity cost of exam preparation for candidates in that bin. The opportunity cost is calculated by multiplying the average forgone wages for candidates in that bin by the number of candidates in that bin.

The main takeaway is that 76% of the total opportunity cost comes from candidates who are not working. This value is large both because a majority of candidates do not work (see Figure 2.3) and because those who do not work forgo more income than those who do (see Figure 2.4a).

Most of the opportunity cost of exam preparation is borne by candidates from poorer households

Table 2.7 repeats the exercise from Figure 2.5, except this time the opportunity cost is split across individuals by their household income.

Table 2.7: Total opportunity cost of exam preparation by candidate household income

Household Income	Opportunity Cost (Cr Rs.)	% of Total
Less than Rs. 5,000	987	18
Between Rs. 5,000 - 10,000	1,780	32
Between Rs. 10,000 - 20,000	1,688	30
Between Rs. 20,000 - 30,000	618	11
More than Rs. 30,000	538	10

Note that about half of the total opportunity cost is borne by the candidates from households earning less than Rs. 10,000 per month.

Even after the exam is over, most candidates who were not working before the exam continue to remain out of work

The months after the last Group 4 exam was a time when it was most opportune for candidates who were not working to take up work. TNPSC will likely not conduct new Group Recruitments until the current ones are processed, and that usually takes time. By their own admission, the average candidate in the Follow-up sample who intends to apply for future TNPSC exams expects the next exam to be about 11 months away.

Even still, among candidates who were not working before the exam, the majority remain unemployed even after the exam:

- Only 42% of these candidates have searched for a job since the exam was over.
- Only 32% have worked in any job since the exam was over

This suggests that shifting candidates from unemployment to employment may be quite difficult for several possible reasons: a) candidates may be reluctant to work; b) candidates may have a hard time finding a job that suits them; or c) candidates may be too invested in the exam process.

How much do candidates study?

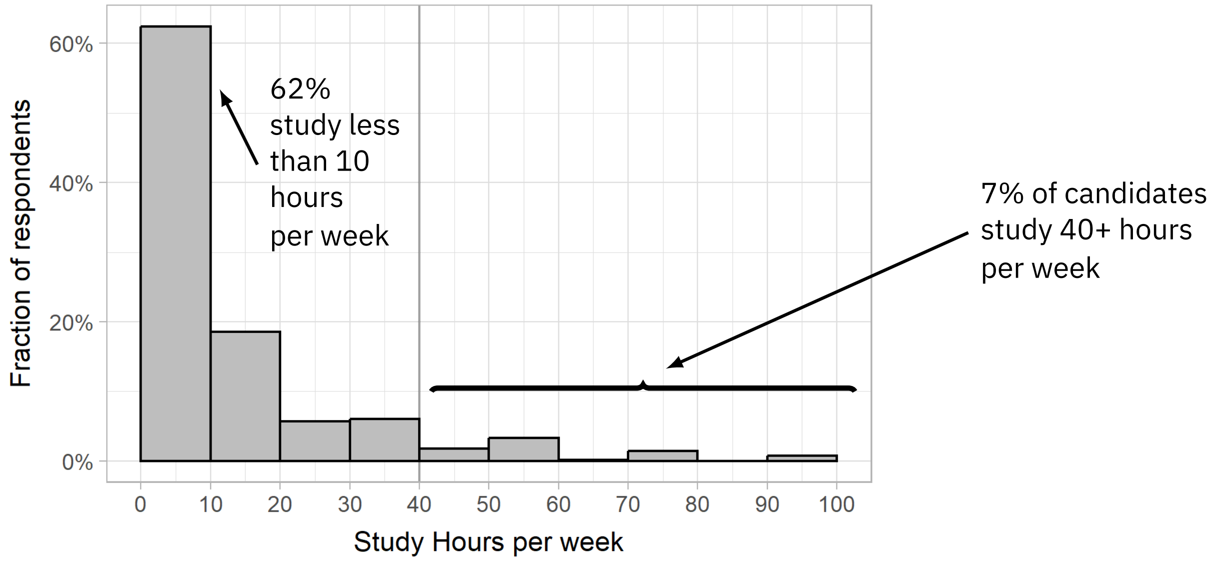
If candidates are not working, then what are they doing instead? Are they studying?

Intense exam preparation is concentrated among a small group of candidates.

We measure study time by asking candidates to report: i) the usual number of days they study per week; and ii) the number of hours they usually study on the days they study. The usual number of weekly hours is the product of these two estimates.

Most candidates do not study much. But there are some candidates who treat exam preparation like a full-time job. (This accords with the distinction that both candidates and coaching center directors made between "serious" candidates and "not-serious" candidates.)

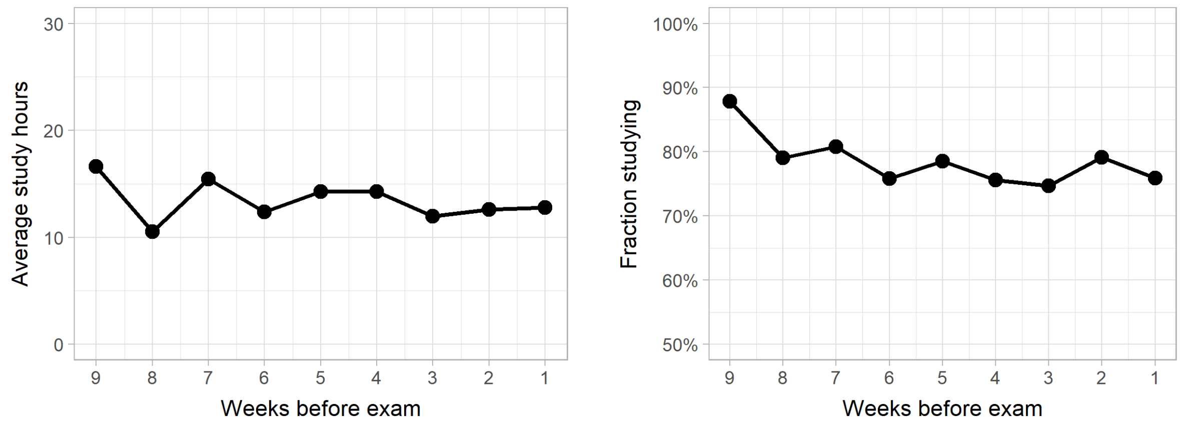
Figure 2.6: How much time do candidates spend studying per week?



There seems to be no special ramp-up of study hours in the weeks leading up to the exam.

About 80% of respondents report spending any time studying. On average, respondents report spending 13 hours per week. This level holds steady in the 2-3 months preceding the exam over which we observe candidates.

Figure 2.7: Study hours relative to exam date



Candidates who are not working do not study much more than those who are

Candidates who are not working only study for about 4.39 more hours on average than those who are working part-time jobs, and 4.72 more hours than those who are working full time jobs.

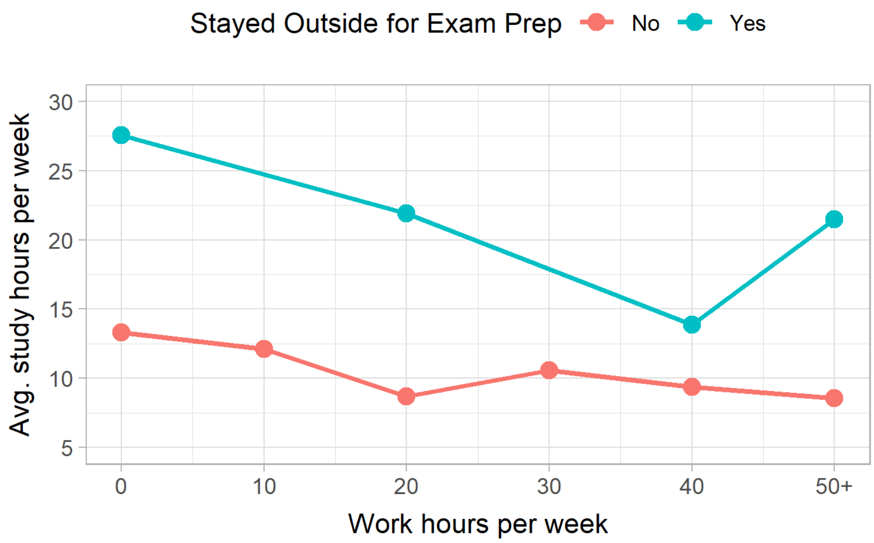
Figure 2.8: The tradeoff between exam preparation and working

There may be a couple reasons why studying does not perfectly substitute for work:

- Some candidates, especially women, may be busy taking care of household chores.
- Candidates may also be engaged in other kinds of formal study, including pursuing undergraduate or master's degrees.
- Some candidates may simply not be very dedicated to exam preparation, and enjoy substantial amounts of leisure time.

The candidates who are most dedicated to studying are those who stay outside of home while preparing (see Figure 2.9). Even those who are working while staying outside study more than those who are at home and not working at all. This suggests that needing time to study is a relatively minor factor explaining why candidates are not working.

Figure 2.9: Candidates who stay outside home while studying spend more time studying

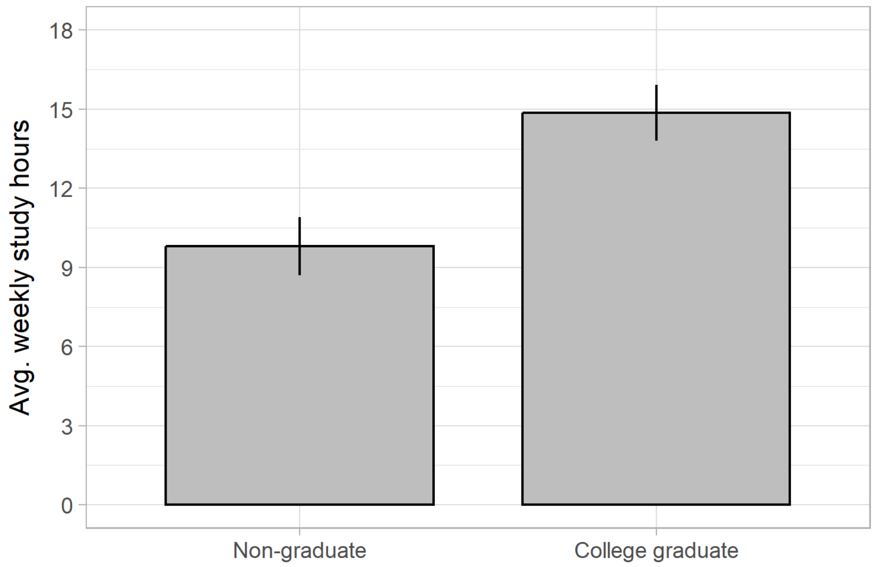


Notes: Cells in which there are few observations have been omitted.

Educated candidates tend to study more

A key consideration for many recruitment agencies is how to make sure that the exam process allows candidates from disadvantaged backgrounds a chance to compete successfully.

Figure 2.10: Average study hours by educational attainment



Even controlling for differences in age, gender, and district, college graduates study 3.7 additional hours per week, or 37% more.

Notes: The black bars indicate the uncertainty in the estimate.

In theory, exam preparation can offer candidates with less education or with gaps in their schooling a chance to catch up with the rest of the competition. This could help level the playing field.

In practice, though, as we see in Figure 2.10, exam preparation only seems to further increase the gaps between candidates. The candidates who study the most are those who have more education. This may either happen because less educated candidates do not have the means to free up time for studying, or because less educated candidates are less inclined to study.

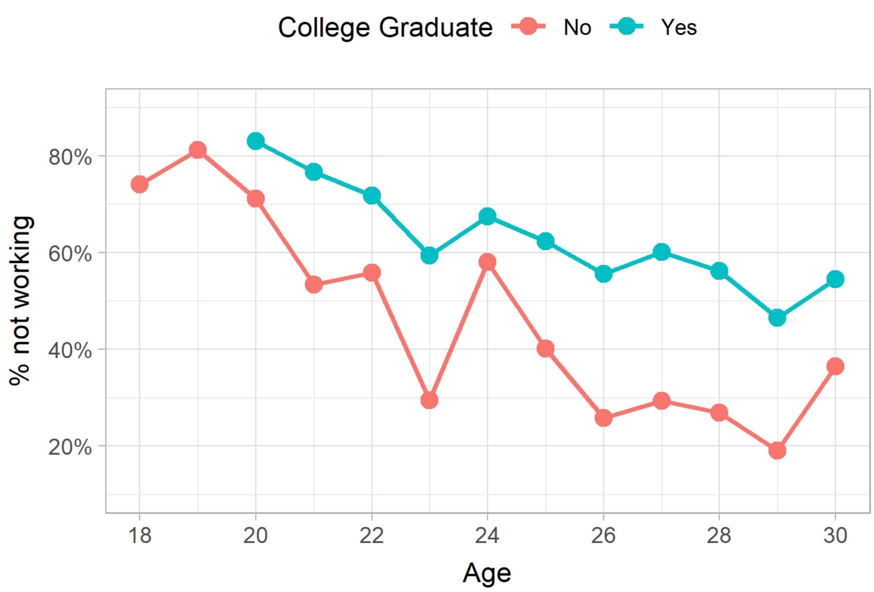
Which candidates are not working?

Who are the types of candidates that are most likely to not be working? Are they not working because they are struggling to find jobs, or are they doing so out of choice?

Young college graduates are least likely to be working

The patterns we see in Figure 2.11 are reminiscent of how application rates vary by age and education (see Figure 1.7). College graduates are much less likely to work than non-graduates, and the fraction of candidates who are not working declines steadily with age.

Figure 2.11: Which candidates are not working? Variation by age and education



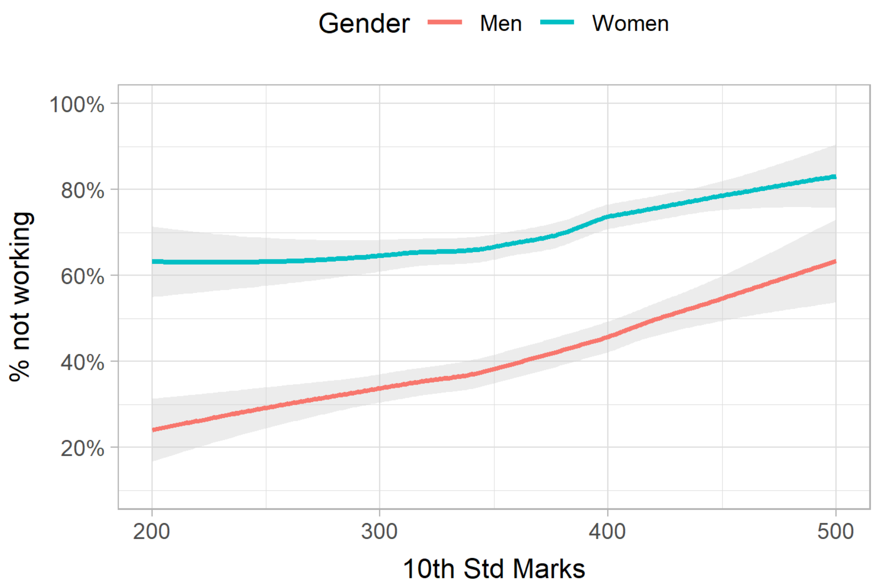
This suggests that there are, in effect, two stages to dropping out of exam preparation: first, decreasing investments in the exam preparation process (what might be called the "intensive margin" of participation); and second, stopping applying for exams altogether (the "extensive margin" of participation).

Candidates who performed better in their school exams are less likely to be working

Figure 2.12 (below) plots the fraction of candidates not working against their 10th standard marks. For both men and women there is a consistent positive slope throughout the distribution.

Since we expect people who did better in school to have better job prospects (all else equal), it is unlikely that candidates who are not working are the ones who are finding it harder to get hired.

Figure 2.12: Candidates with better school performance are more likely to be not working



Notes: Each line plots a local linear regression. The grey bands capture the uncertainty in the estimate. For clarity, candidates with less than 200 marks on their 10th standard board exam were dropped. These account for less than 1% of observations.

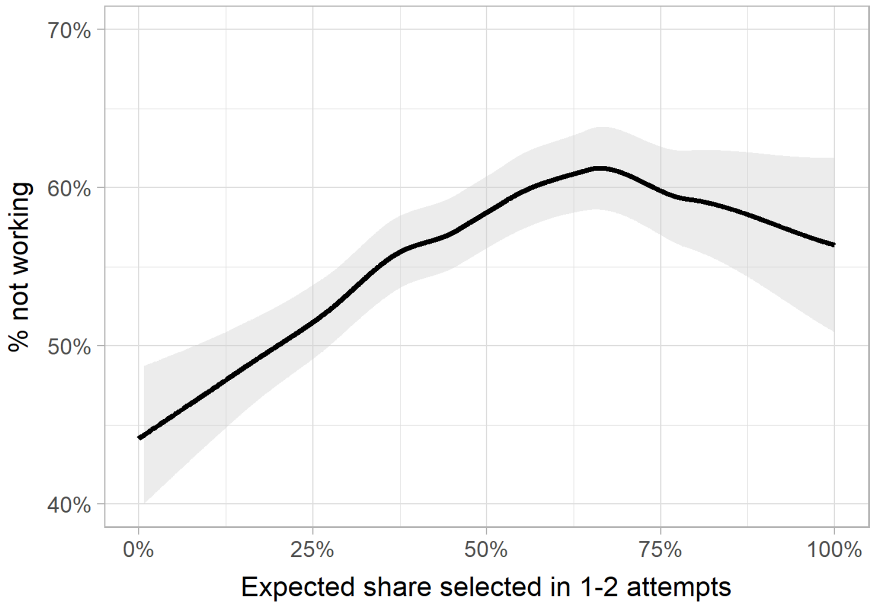
Candidates who think it is easier to get selected are less likely to work

In the Candidate Survey, we asked candidates to guess what fraction of selected candidates in the 2019 Group 4 needed

3 or more attempts to get selected. By subtracting this response from 100, we get a measure of the share of candidates they think are selected in 1 or 2 attempts. This is, in a way, a proxy measure for how difficult they think the exam is. Candidates who think it is harder to get selected will, naturally, believe that very few candidates selected in less than 3 attempts.

We find that candidates who think the exam is easier are also those who are less likely to work (see Figure 2.13). The difference in employment rates between the most and least optimistic candidates is about 10-15 percentage points.

Figure 2.13: Candidates who believe the exam is easier are more likely to be not working



Notes: Each line plots a local linear regression. The grey bands captures the 95% confidence interval of the main estimate.

Summing Up

Most candidates are not working, but that does not mean they spend their full time studying. Intense studying is concentrated among a small group of candidates who stay outside of their home. The reasons why candidates remain out of work appears to have less to do with the demands of exam preparation, and more to do with their academic background, and their optimism about getting selected soon.

2.4 Exam preparation as a household decision

Who decides how much to invest in exam preparation? Do candidates decide on their own? How much influence do parents or relatives have?

Candidates rely on external support to finance their investments in exam preparation

In the Candidate Survey, we asked respondents who reported making any expenditure on exam preparation who they relied on for financial support.

About half of the sample reports relying on others for financial support in some capacity. Although in the survey we did not probe whether this financial support came from family, my conversations with candidates suggest that this is the most likely source.

Table 2.8: Sources of funding for candidates who made some direct expenditure on exam preparation

Funding Source	Overall Share (%)	Exclusive Share (%)
Family and Friends	50	42
Current Income	34	23
Savings	30	18

Notes: Candidates could select multiple categories, so columns can total to more than 100%. "Overall Share" refers to the share of respondents who take funds from the given source. "Exclusive Share" refers to the share of respondents who rely exclusively on the given source.

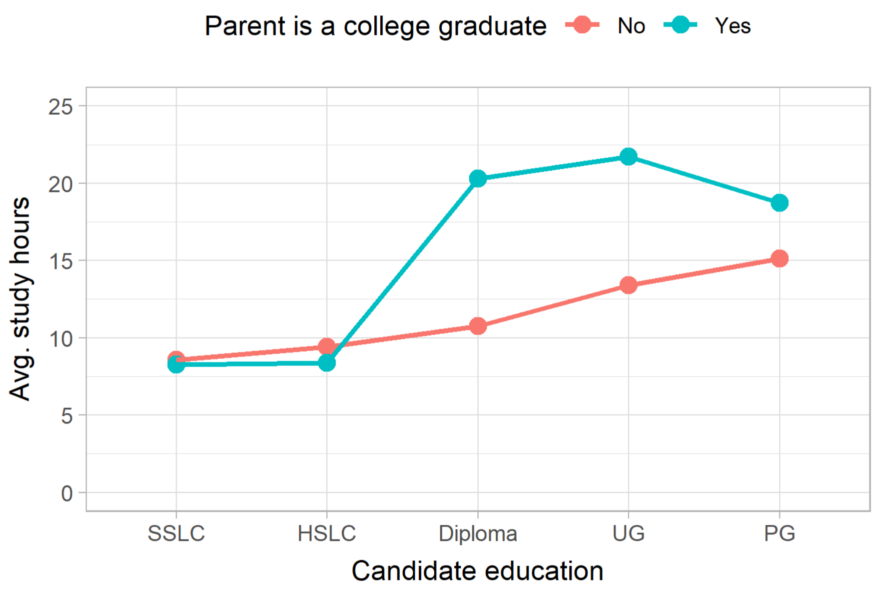
Candidates who use funds from family and friends tend to rely exclusively on them. As a result, families can have influence over candidates' exam preparation decisions. For example, they may put a time limit on the amount of time they are willing to continue to provide financial support.

Candidates who come from more educated families invest more in exam preparation

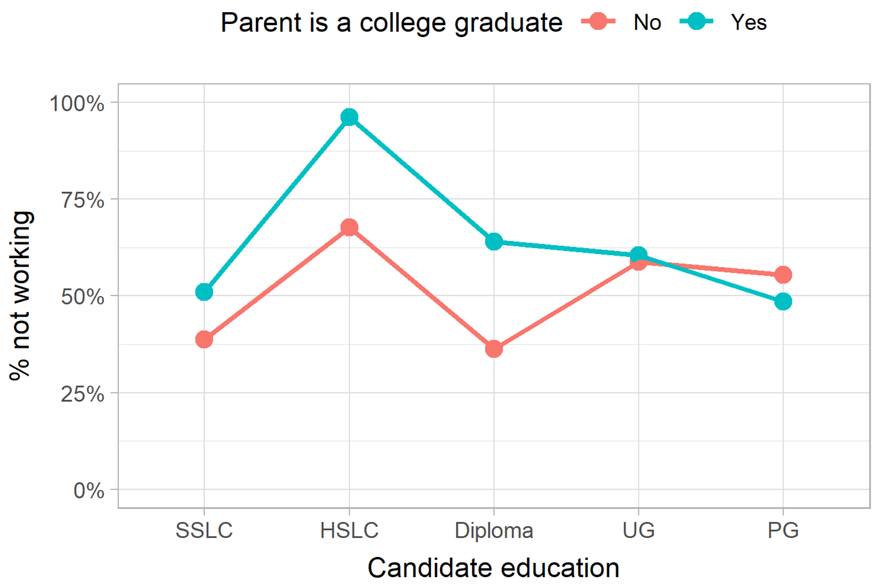
So far we have considered three different measures of candidates' investment in the recruitment process: time spent studying, not working, and direct expenditure. Across all three measures, candidates whose parents are more educated invest more—regardless of how much education candidates have themselves.

Figure 2.14: Investment in TNPSC exam preparation by candidate and parental education

(a) Hours studying per week



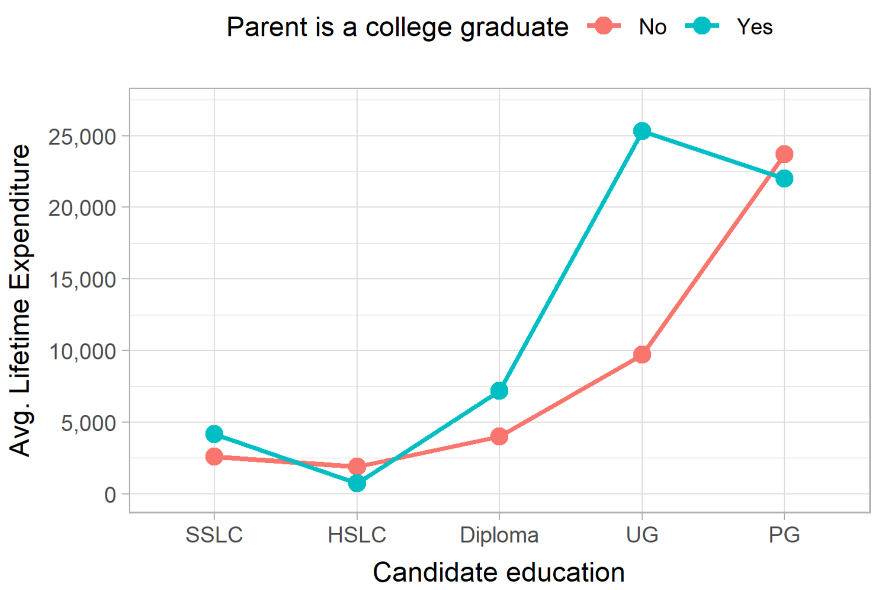
(b) Not working



Notes: Figure continued on the next page.

Figure 2.14 (cont.): Investment in TNPSC exam preparation by candidate and parental education

(c) Lifetime direct expenditure on exam preparation



Notes: Abbreviations: SSLC = 10th Standard; HSLC = 12th Standard; UG = undergraduate degree; PG = post-graduate degree.

In our interviews with candidates, those who had relatively less educated parents reported unique challenges in winning support from their families for their studies. In particular, these candidates often reported that their parents did not understand how difficult it was to clear the exam. Instead, their parents generally expected that any serious aspirant should be able to get selected on their first attempt.

The reason why uneducated parents' expectations were so misaligned is that they understood very little about the exam process. Some parents felt shy or even embarrassed by their own lack of education to ask questions of their children to learn more. But in some cases this lack of information was strategic. Candidates often anticipated that their parents would not allow them to study if they knew just how risky or how long it would take to get selected. Thus, to buy themselves time, candidates would employ a range of strategies for managing their parents' expectations. For example, rather than outlining their true intentions, in each attempt, they would say, "it will happen this time," and then count on their parents continuing to give them permission each time the topic came up.

In future research, it is important to develop a better understanding of who is the primary decision maker regarding exam participation and investment—candidates or their family members.[31] On the one hand, candidates typically have more information; but on the other hand, parents can control candidates' access to funds. Which of these factors is more important is an empirical question.

[31] One of the key challenges of answering this question is talking to parents. Given how carefully candidates manage communication with their parents, in our experience it was common for candidates to refuse us permission to talk to their parents directly, worried that we might destroy their parents' support for their exam preparation.

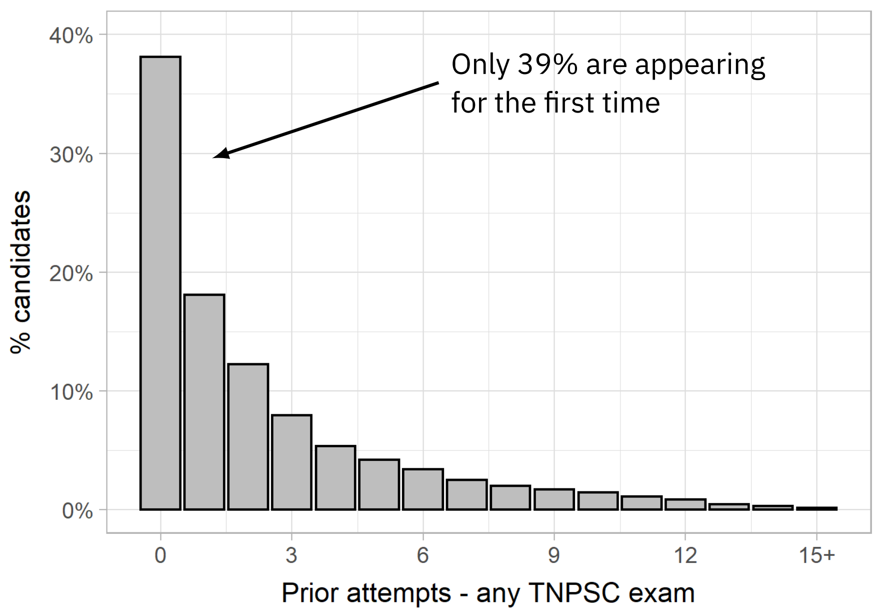
2.5 Persistence

So far we have looked at the cost of exam preparation *per recruitment*. What do these costs look like from the candidates' point of view, over the course of their careers?

Most candidates appearing for the exam have made several prior attempts.

Figure 2.15 looks at the number of prior attempts that candidates have made in any TNPSC exam. The sample is restricted to recruitments at the end of our sample window (i.e. FY 2018/2019) so that we can observe the longest possible window. The figure plots the distribution of the number of appearances in TNPSC exams that were notified between FY 2013 and FY 2017.

Figure 2.15: Number of previous attempts in any TNPSC exam among candidates appearing in FY 2018/2019

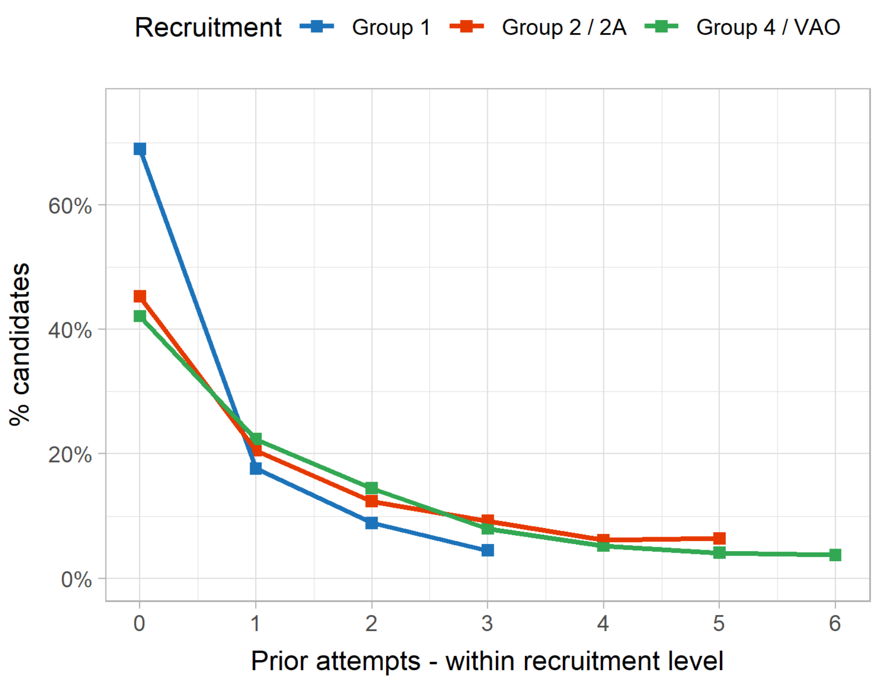


Most candidates are not appearing for the first time. However, extended test-taking is not common either. Although there are no limits on the number of exams candidates can take, only 10% of candidates appear for more than 6 attempts, which is the most strict limit on number of attempts in place for civil services exam (namely, this is the limit on the number of attempts in the general category for the UPSC exam).

Candidates drop out of the Group 1 exam faster

Figure 2.16 plots the number of prior attempts separately by exam group. For example, the Group 1 line plots the distribution of number of prior attempts specifically in Group 1 exams for candidates who appeared in the 2019 Group 1 exam.

Figure 2.16: Number of previous attempts in any TNPSC exams by group



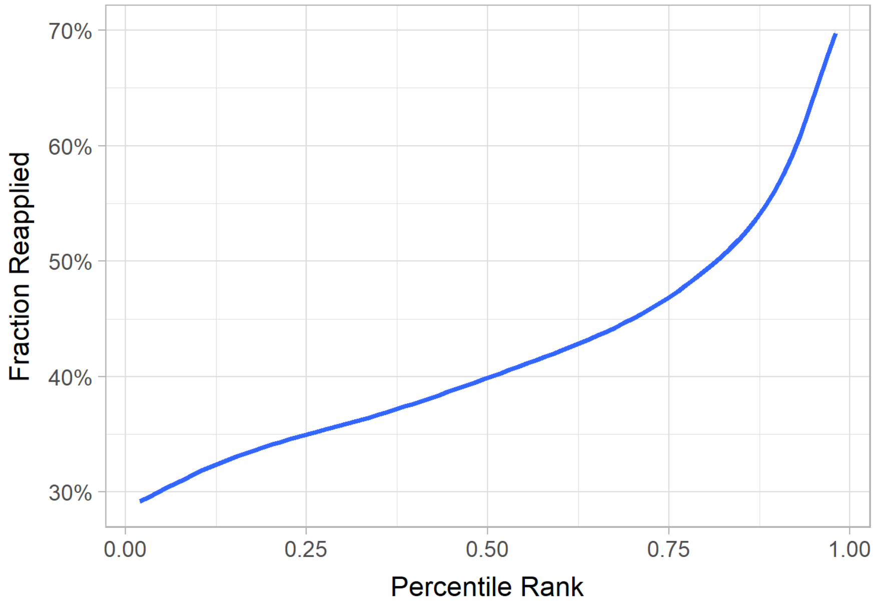
Notes: The Group 1 line plots the distribution of the number of prior attempts in Group 1 exams for candidates appearing for the 2019 Group 1 exam. The Group 2/2A line does the same for candidates appearing in the 2018 Group 2 exam. The Group 4 line plots the same for candidates appearing in the 2019 Group 4 exam.

The number of first-time candidates in Group 1 is about 50% larger.

Candidates with higher marks are more likely to re-apply

How does the composition of candidates change as they make more attempts? Do the repeat test-takers tend to be the candidates with higher scores or the ones with lower scores?

Figure 2.17: Re-application rates across Group 4 exams given previous exam score marks



Notes: The figure shows the probability of re-applying in the next available Group 4 based on candidates' prior exam marks. The exams included in the sample include: 2013/09, 2014/18, 2015/19, 2016/15, 2017/23, and 2019/19.

In Figure 2.17, we see that candidates with the highest scores are more than twice as likely to apply as candidates with the lowest scores. Interestingly, about half of the variation in average re-application rates is determined in the top 10 percent of the score distribution, and the rest is spread out over the remaining 90 percent.

I do not yet know whether candidates who invest more also tend to score well. But if they do, that means that the candidates who invest the per year are also more likely to invest *across* years. As a result, the variation in investments we see across attempts will be even larger than the variation within a single attempt that we have seen so far.

Are these investments worth it?

Candidates invest substantial resources into exam preparation. Are these investments worth it? As it turns out, this question cannot be answered directly. That is because we don't know the value of a government job to begin with.

The wage is just one component of what makes a government job valuable. Government jobs also come with many other *amenities*, such as job security; status and respect; preference in the marriage market; and in some cases, access to bribe payments.

Given all these additional benefits, it is very likely that candidates value government jobs more than the salary amount. But how much? In order to assign a monetary value to these amenities, we need to know what kind of salary would induce candidates to give up a government job in exchange for a private sector job without these amenities. But answering this question is challenging, because it is extremely rare (if not unheard of) to observe a candidate making this kind of trade-off.

We could, however, ask the question in another way: How much would a government job need to be worth in order for candidates' investments to be worthwhile? To answer this question, we solve for the value of government jobs that would ensure that candidate's investments have positive return. This gives us an indication of how much the rationality of candidates' investment depends on their subjective (and potentially malleable) assessment of the value of a government job.

In Mangal (2022a), I estimate the amenity value of a government job in the context of candidates preparing for the Maharashtra Public Service Commission exam in Pune. In a sample of highly motivated candidates, I find that government jobs are likely worth at least *Rs. 4 lakh per month*. This is at least 5 times the amount that candidates would earn just from the base salary alone.

Thus, even though India has one of the largest public-private wage differentials in the world (see Finan et al. (2017)), the salary gap alone is not enough to explain why highly motivated candidates invest so much in exam preparation.

For additional details, see Mangal (2022a). "How much are government jobs in developing countries worth?."

2.6 Discussion

Is there too much investment in exam preparation?

Who benefits from candidates making intense investments in exam preparation? Does it just help some candidates get posts at the expense of another equally qualified candidate? Or do these investments generate some broader social value?

Tournaments with a fixed number of vacancies—like public sector recruitments–are particularly vulnerable to suffering from the problem of too much competition. In the private sector, when there are more workers who are looking to supply a given set of skills, firms' recruitment costs for that role decrease. Over time, this causes firms to post more vacancies demanding those skills, which ameliorates some of the congestion in the labor market.[32]

[32] For example: The growth of the IT industry in the US encouraged many Indian students to learn computer science. Because the number of US visas were capped, many of these students remained in India. The ready supply of computer skills in India fueled the Indian IT tech boom, which ended up surpassing the US in IT exports. For more details see: Khanna and Morales (2021).

In competitive exams, there is no such offsetting effect. If candidates choose to put extra effort into studying, the total number of vacancies does not expand. Instead, it puts pressure on the remaining candidates to compete more to keep up. This logic suggests that, in the absence of a braking force, there is a natural tendency for merit-based exams to reach inefficiently high levels of competition.

Should the government consider tightening limits on the number of attempts candidates can make?

Over the past few decades, state governments have weakened age limits and limits on exam attempts at Public Service Commissions around the country. One way of mitigating the effects of excess competition is to tighten these rules again.

If candidates want to invest heavily in exam preparation then that is their prerogative. The problem is that candidates who study more also force others to invest more in order to keep up. These "rat race" effects can lead to an ever-increasing spiral of investment in exam preparation, *even when all candidates would prefer to invest less*.[33] Seen in this light, limits on exam participation could, potentially, make the majority of candidates better off by making it less onerous for them to stay competitive in the recruitment process.

[33] This dynamic is known in game theory as the Prisoner's Dilemma, and provides motivation for many types of regulation that can be welfare-enhancing because people do not fully internalize the costs of their actions (e.g. pollution regulation, quality standards, and more).

A stand-alone policy that tightens limits on exam participation may not be politically feasible. Any new policy in this area will need to contend with the reasons why these limits were loosened in the first place:

- Extreme delays in the process of conducting recruitments has meant that candidates often age out before they have a chance at making several serious attempts. This has meant that recruitment agencies often raise the age limit as a way of compensating candidates for this loss.
- Limits are often seen as regressive. Critics of limits argue that candidates from poorer backgrounds need extra time to catch up, either because they tend to learn about competitive exams later in their career, or because they face unique challenges while studying.

One possible way forward is to combine reforms on participation limits with other reforms that reduce the likelihood of exam delays, and increase poor candidates' access to information, coaching, and support. Moreover, reforms that limit participation could be designed so that they are conditional on candidates' past performance. Not being allowed to reappear is most upsetting for someone who has just missed the cutoff. But it may be easier for recruitment agencies to publicly justify the limits for cases where candidates have repeatedly missed the cutoff by a wide margin. A "bargain deal" of this sort could potentially make a meaningful dent in the social cost of recruitment.

Are there ways of increasing the value of exam preparation for candidates who are not ultimately selected?

The other way of mitigating the negative side effects of competition is finding ways of increasing the value of exam preparation outside of the public sector.

For example, it is worth investigating the value of issuing certificates of exam performance that candidates could use in the private sector. Along these lines, the Union Public Service Commission creates a database on not-selected candidates' exam performance for those who choose to opt into the list. The hope is that this kind of disclosure can improve not-selected candidates' employment outcomes.

However, there has been no systematic study done of the effect of these kinds of policies. Previous research has shown that providing jobseekers with skill certificates can be an effective way of improving search outcomes (Carranza et al., 2022); but it is unknown whether employers value any of the skills tested in civil service exams. This question merits further investigation.

3 Who gets selected?

3.1 Motivation

What are the risks that candidates take in the selection process?

On average, participating in a TNPSC Group recruitment looks like a very risky gamble. Across all exams conducted between FY 2013 and FY 2019, the average selection rate was never less than 1 in 100, and could be higher than 1 in 1,500.

Table 3.1: Applicant to vacancy ratio in TNPSC exams, FY 2013 - FY 2019

Exam Level	Average Ratio	Min. Ratio	Max Ratio
G1	1261	893	1635
G2	321	151	444
G4	320	141	949

Notes: The ratios presented in this table are calculated as the number of candidates appearing for the exam divided by the number finally selected.

But the average selection rate does not necessarily tell us any *individual* candidate's experience. The risk of exam preparation depends on how the individual probability of success is distributed across the population. For example, consider an exam in which 10% of candidates have close to a 100% chance of selection, and 90% of candidates have close to a 0% chance of selection. In this scenario, the average selection rate is about 10%—even though no individual candidate has a selection rate close to this value.

The variation in selection odds across the population sets incentives for candidates.

In Chapter 2, we saw that some candidates invest heavily in exam preparation. How much of this is a response to the incentives built into the design of the test? Does the exam reward persistence?

Are candidates well-informed of the risk they are taking?

The answer could go either way. On the one hand, candidates tend to know other candidates.[34] By sharing experiences, candidates may be able to make reasonable inferences about what it takes to get selected.

[34] In the Candidate Survey, 42% report belonging to study groups, and 35% report knowing someone personally who was selected through TNPSC.

On the other hand, it is common for jobseekers to not be fully informed about the uncertainties of the labor market. For example, the unemployed tend to be over-optimistic about their chance of finding a job, or how long their unemployment spell will last; and they tend to under-estimate the value of adopting new job search strategies that will improve their chances of finding a job (Altmann et al., 2018; Mueller et al., 2021).

The key question is whether candidates can eventually learn about these uncertainties *accurately* and *quickly*. Is there a robust feedback mechanism that helps candidates learn from their experience?

3.2 Variation in selection rates

Not every candidate has the same probability of selection. How much heterogeneity is there?

The probability of selection varies dramatically across candidates

To uncover the variation in the probability of selection, I use a machine learning algorithm called *random forests*. For more details about the estimation strategy, and well the model fits the data, see Appendix B.1.[35]

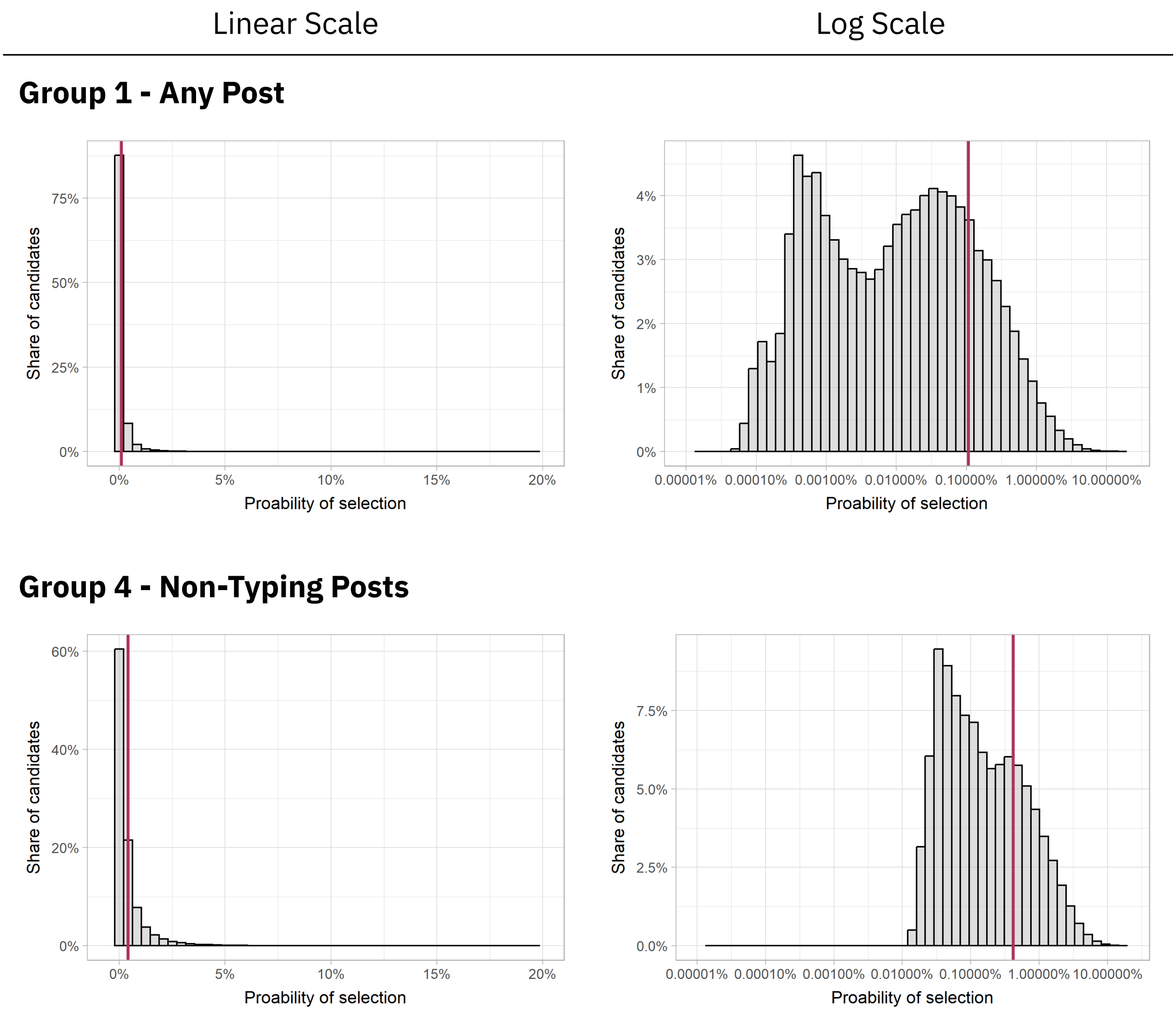
Figure 3.1 plots the estimated distributions for the 2019 Group 1 and Group 4 exams. The maroon line corresponds to the average selection rate. Because the probability of selection varies so much, I present the distribution on both a linear scale, and a log scale. The former shows just how skewed the distribution is. The latter allows us to zoom in on the variation in selection rates for the bottom 99% of the population.

For the Group 4 exam, the sample is restricted to candidates who do not qualify for typing or steno-typist posts. Less than 7% the candidate pool qualifies for these positions, so the competition for these posts functions like a parallel contest within the overall Group 4 contest. Ignoring people applying for these positions allows us to focus on heterogeneity in candidate characteristics rather than vacancy availability.[36]

[35] The main takeaway is that the model is able to predict the average selection rate within 5% of the true value. The fit is best in the middle of the distribution, and deteriorates for the bottom and top 1-2% of predicted values.

[36] In the 2019 Group 4 exam, the probability of selection for typists and steno-typists was 5%, which is 11 times the rate of non-typists. These posts are much less competitive in part because they require expensive certifications, which imposes a barrier to entry.

Figure 3.1: The distribution of the probability of selection



Notes: The figure plots the distribution of the predicted probability of selection across candidates. Each column plots a different view of the same underlying distribution. The distributions are estimated through a machine learning algorithm, which generates a predicted probability of selection based on covariates. Details on the estimation procedure can be found in Appendix B.1.

Table 3.2 summarizes the distribution. There are three features that stand out:

- The probability of selection varies dramatically across candidates. For example, in Group 4, the top 5% of candidates have a selection probability that is close to 100 times higher than that of the bottom 5%. In Group 1, that differential is over 2,500 times.
- Despite the large variation, for most candidates the probability of selection is still very low. Even at the 95th percentile, the selection rates are still no higher than 2%

in Group 4.

- As low as the average selection probability is, for 75% of candidates their selection probability is even lower.

Table 3.2: Quantiles of the distribution of the probability of selection

Percentile	Group 1 (%)	Group 4 (%)
5	0.00019	0.02337
10	0.00034	0.02946
25	0.00094	0.04721
50	0.01065	0.12508
75	0.07038	0.43617
90	0.26661	1.07065
95	0.52272	1.72744
Average	0.10739	0.42028

Candidates who register for the exam at the last minute are less likely to get selected

Why do such a large share of candidates have a low selection rate? One important reason appears to be that many candidates who apply are not very serious.

One way we can measure how serious candidates are is by looking at when they register for the exam. Recall, candidates have about a month to register for exams when they are first notified. The exact date on which registration will open is generally unknown. As a result, the candidates who will be among the first to find out when registration opens are those who are paying close attention. Serious candidates also tend to prefer to register as soon as possible, in part to have a better chance of getting allotted their preferred exam hall location. Candidates who are less serious tend to register at the last minute.

Figure 3.2 shows just how much the probability of selection can vary along this axis.

In Group 4 exams, candidates who register on the first day have a probability of selection that is 27 times higher than candidates who register on the last five days.[37] In Group 1, the differential is not as large—about 2 times—but we do see

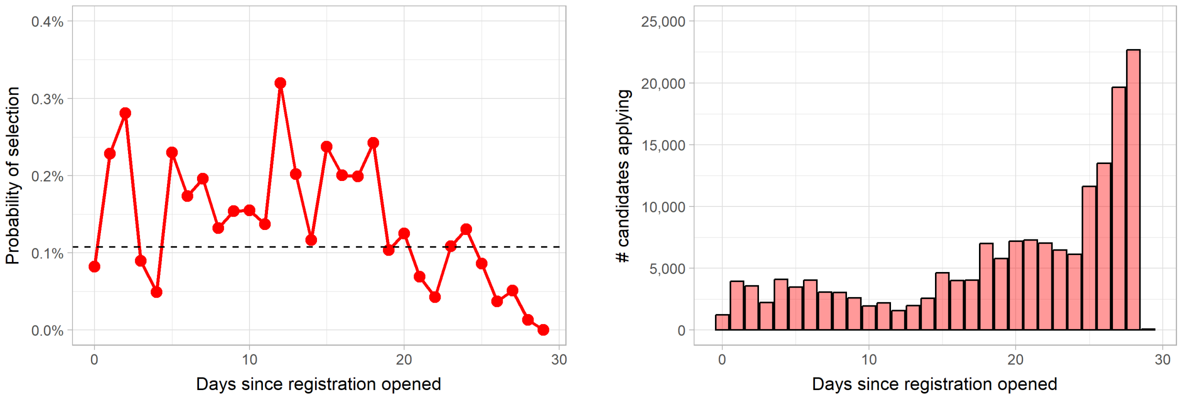
[37] This is a large effect, equivalent to moving from the 25th percentile of the distribution of selection probabilities to the 90th percentile (see Table 3.2).

the same drop off in selection probabilities for candidates that register at the last minute.

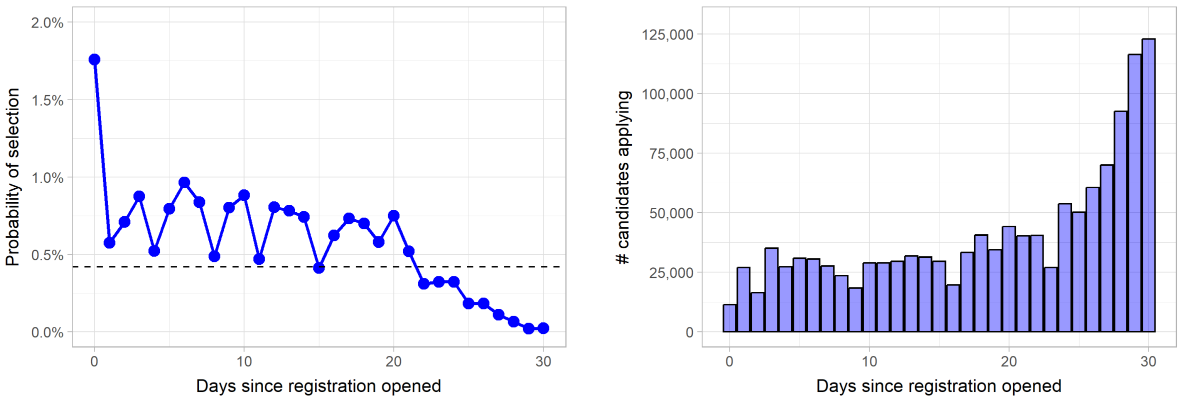
Figure 3.2: Candidates who register late have lower selection probabilities

Probability of Selection	Registration Timing

Group 1 - Any Post



Group 4 - Non-Typing



Notes: The dashed line in the first column marks the average selection rate in the population. The probability of selection is the share of candidates who applied that day who were ultimately selected.

Incidentally, the last week is also when a large share of candidates register. About 36% of Group 4 candidates and 40% of Group 1 candidates registered in the last five days.

In my conversations with candidates, I learned that many of these last minute registrations are driven by a kind of 'follow-the-crowd" mentality—especially in Group 4. Those who haven't already applied will be encouraged (or perhaps even pressured) to do so. Even though they haven't been studying, or perhaps are not very confident, they figure they might as well give it a shot. In the worst case scenario, they lose an afternoon, and in the best case, they are set for life![38] I suspect that the reason we do not see such a stark slope in Group 1 is because participating in that exam is not something people take as lightly.

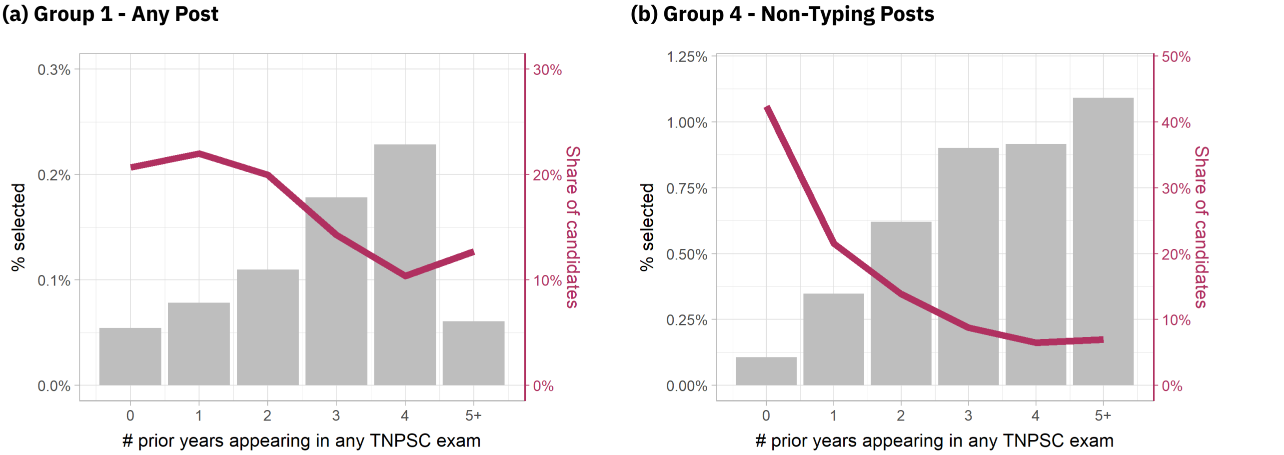
[38] Before the 2022 Group 4 exam notification was released, I visited a village in Kancheepuram district to talk to some candidates. *There are twenty of us studying now*, one of the candidates in the village said; *but just you watch, about two hundred people from this village will apply*.

Candidates with more prior exam experience have a higher chance of selection.

The other reason why candidates are not equally competitive is that they have very different levels of prior experience. Inasmuch as preparing for TNPSC exams is an education of its own, there is no reason to expect that first-time candidates are able to compete with their more experienced counterparts.

Figure 3.3 shows that prior experience matters. Candidates who have appeared for several prior attempts have a probability of selection that is 4 to 5 times that of candidates making their first attempt. At the same time, the vast majority of candidates have made few prior attempts. As a result, a large share of the applicant pool has lower-than-average selection rates.

Figure 3.3: Probability of selection and prior experience



Notes: The horizontal axis measures the total number of fiscal years in which the candidate appeared for any TNPSC exam, from FY 2013 to FY 2017. (FY 2018 is excluded because it had a single Group 2 recruitment.) The maroon line marks the share of candidates belonging to each bin along the horizontal axis (see right axis). The grey bar measures the share of candidates that were selected (see left axis).

3.3 How much does luck matter?

There is some component of the underlying risk of exam-taking that is not in the candidate's control. We'll call that component *luck*.

In single-stage exams, luck arises from the fact that, even for questions drawn from the same syllabus, candidates may happen to know the answer to some and not to others.[39] As a result, candidate's final score can fluctuate.

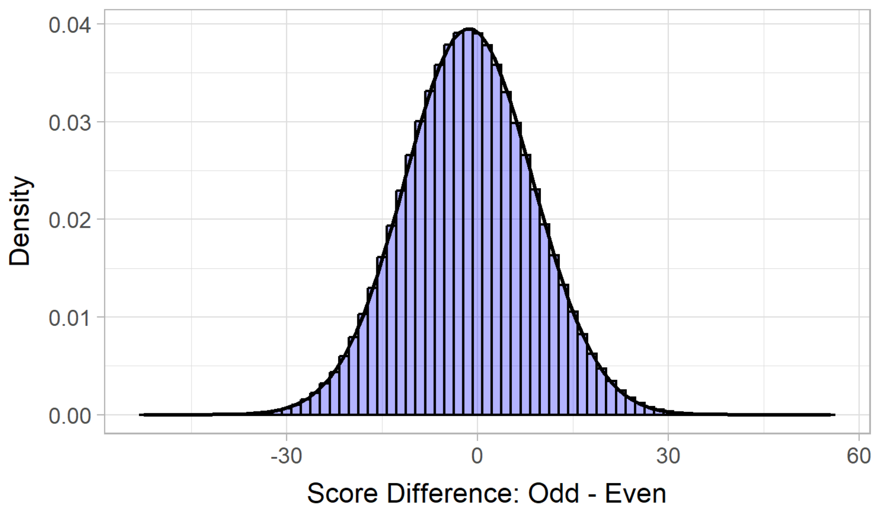
[39] In multi-stage exams, luck also depends on who is grading the written exam and the interview.

How much does this component of risk matter? In this section, we look at this question in the context of Group 4 exams.

In Group 4 exams, about 7-12% of the overall variation in exam scores is due to luck.

To separate luck and ability, I measure the difference in test scores between the odd and the even questions on the exam. This controls for all factors that are fixed on the day of the exam, such as test conditions, and candidates' exam preparation. The only remaining factor that explains the difference in scores is the part that we want to measure—namely, that candidates get lucky with some questions (either because they know the answer, or make a good guess) and not with others.

Figure 3.4: Variation in test scores due to luck, Group 4



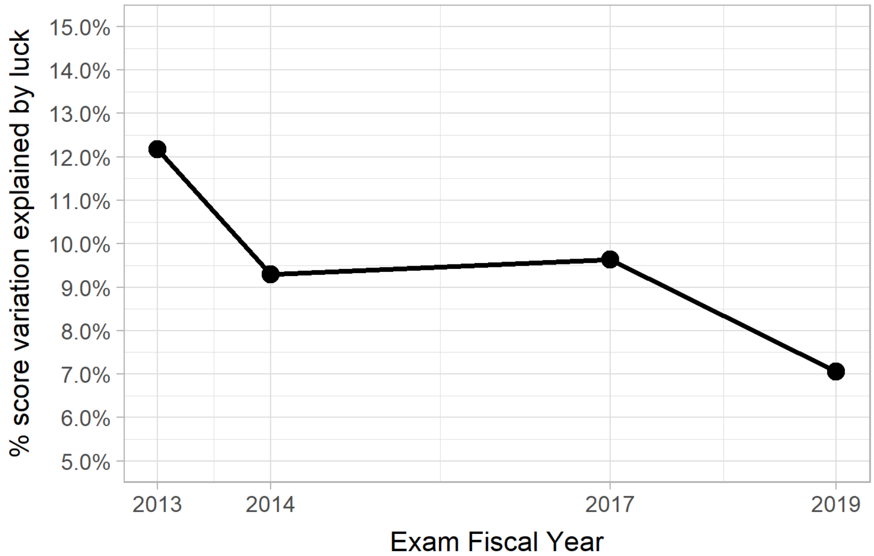
Note: The figure is based on data from the 2013/09, 2014/18, 2017/23 and 2019/19 Group 4 exams. The blue bars are the empirical histogram. The solid black line is the normal density with the same mean and standard deviation.

Figure 3.4 shows that the difference in test scores between the odd and even questions on the exam follows a normal distribution.

The fact that this variation follows allows a normal distribution allows us to back out what share of the total variation in exam scores in Group 4 is due to luck.[40] Figure 3.5 plots the share of total variation that is explained by luck in four prior Group 4 exams. Interestingly, there has been a downward trend over this period.

[40] See Appendix B.2 for a discussion of the assumptions required for this calculation.

Figure 3.5: The share of variation in test scores explained by luck, Group 4



Note: The figure is based on data from the 2013/09, 2014/18, 2017/23 and 2019/19 Group 4 exams. See Appendix B.2 for details on how the vertical axis was calculated.

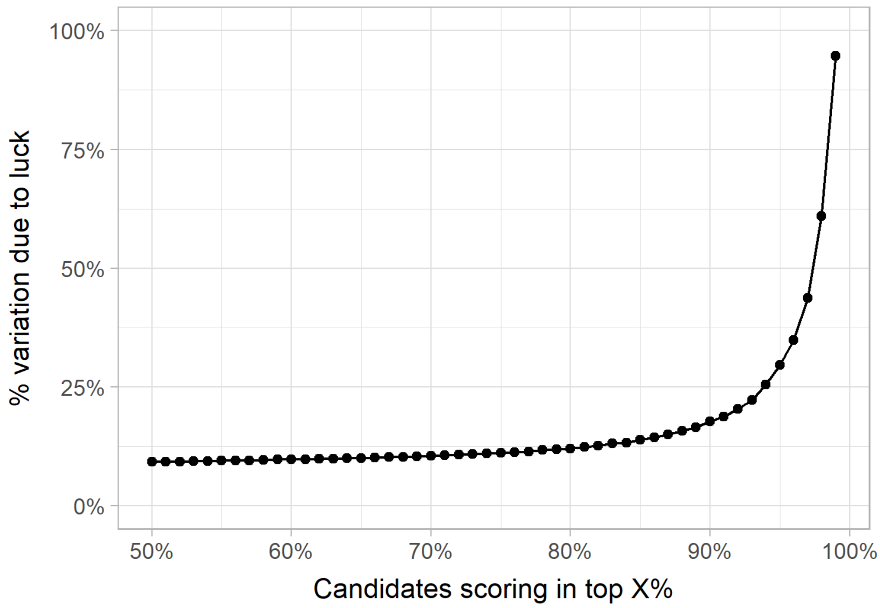
The variation in test scores due to luck is highest at the top of the distribution

The gap in test scores between candidates is not even across the distribution. As a result, the share of total score variation explained by luck is not constant across the distribution either.

Figure 3.6 shows how the share of total variation explained by luck depends on how finely we want to distinguish between candidates. I use the 2019 Group 4 exam as an example. The data point at the far left end of the graph plots the share of total variation in the top 50% explained by luck. In other words, if we want to figure out whether a candidate has average ability or very high ability, the test is able to do so fairly reliably.

On the other hand, at the far right end, the graph plots the same statistic for the top 1% of candidates. Here, almost all of the variation in scores is explained by luck. In other words, the test cannot reliably distinguish someone whose ability is at the 99th percentile from the candidate whose ability is at the 99.9th percentile.

Figure 3.6: Variation in test scores explained by luck for top-performing candidates



Notes: For each data point, the sample is restricted to candidates whose ranks are equal to or larger than that value, and then the share of variation explained by luck is calculated as explained in Appendix B.2. Each point marks a difference of one percentile.

This is a natural consequence of having a test with a finite amount of precision and a very large number of applicants.[41] At the top of the distribution, the variation in ability is small, while the limitations on the precision of the test stay relatively constant. As a result, most of the variation at the top of the distribution is due to luck.

[41] In Chapter 5, I discuss in detail how the precision of exams can be increased.

For candidates who are on the margin of selection, this can be frustrating, because it implies that whether or not they are selected in a given attempt is largely out of their control.

Figure 3.7 illustrates with an example from the 2019 Group 4 exam. In this exam, about 81% of all candidates who were selected in non-typing posts needed a score in the 99.5th percentile or above in order to get selected.[42] Let's see how much a small variation in scores due to luck can affect whether

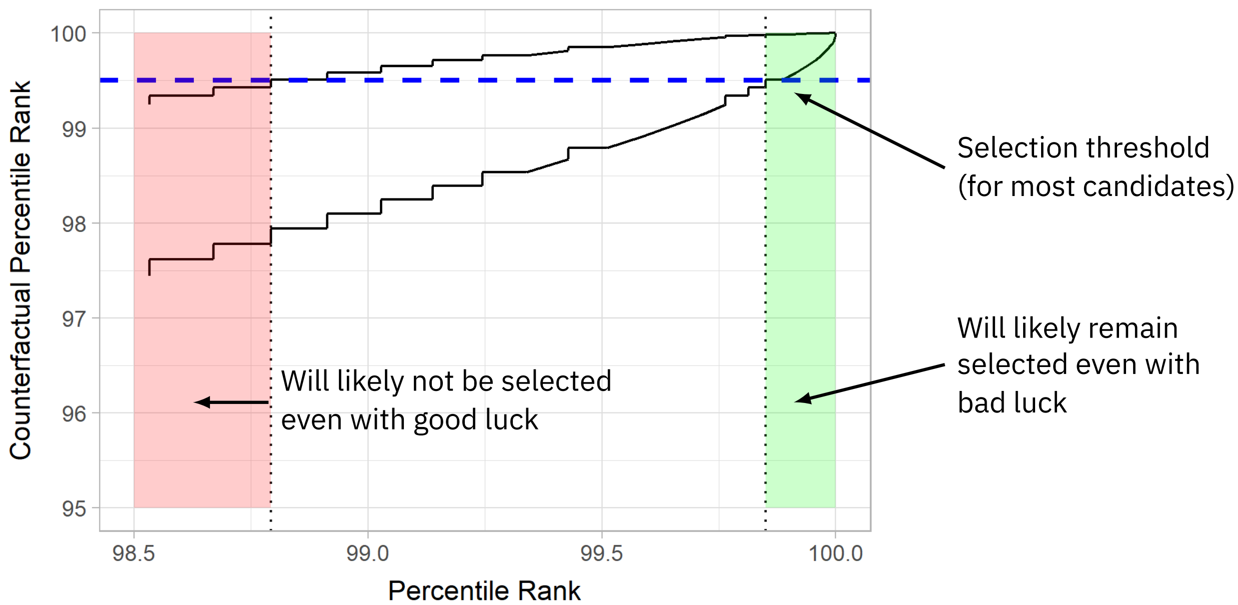
[42] The remaining 19% were selected with lower scores because of TNPSC's various reservation policies.

candidates are able to cross that threshold.

In Figure 3.7, the black lines show how much candidates' rank can vary with a one standard deviation difference in luck. This means that, for candidates whose "true" rank is given by the horizontal axis, luck will cause candidate's rank to land somewhere between the black lines 68% of the time.

The blue dashed line marks the selection threshold at the 99.5th percentile. If it lands in between the black lines, then that candidate will sometimes get selected, and sometimes not. The candidates on the far right, marked in green, will get selected even if their scores change a bit due to luck. The candidates on the far left, marked in red, will not get selected, even if luck boosts their scores by a bit.

Figure 3.7: How sensitive is selection to luck?



Notes: Graph uses data from the 2019 Group 4 exam. See main text for an explanation of the figure.

We see that candidates as far down as the 99th percentile have a reasonable chance of making it into the top 0.5%. In other words, about twice as many candidates are competitive for selection as the exam intends to select. The good news is that this is not a large margin overall. Only about 10,000 applicants need to worry about luck affecting their chance of selection, out of a population of over 13 lakh.

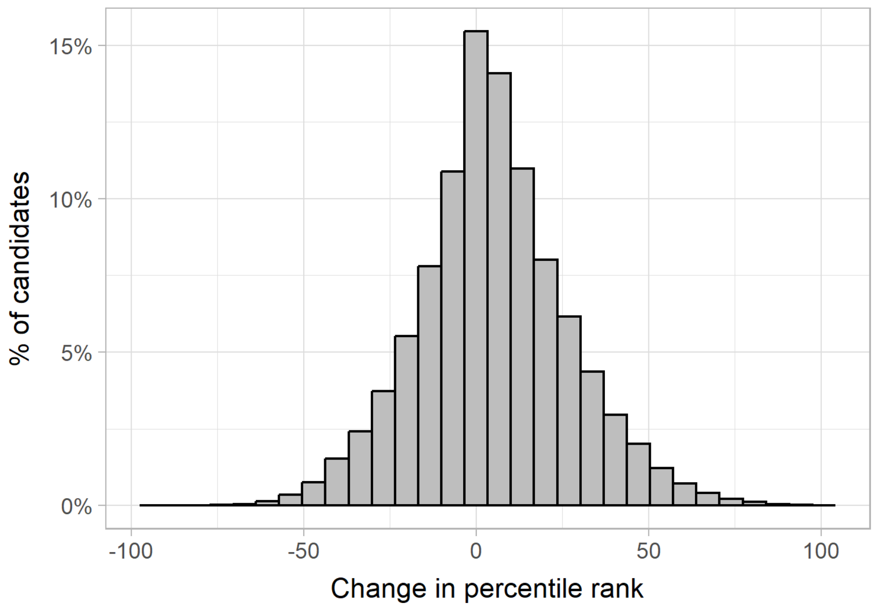
3.4 The value of additional attempts

How much does it pay off to make additional attempts? Does candidates' ability continue to increase as they make more attempts? Or do they tend to stagnate without fundamentally increasing their chances of selection? These are the questions I address in this section, focusing on the case of Group 4 candidates.

On average, exam performance does not improve with additional attempts.

Figure 3.8 plots the change in candidates' performance in subsequent attempts in Group 4 exams. For most candidates, the change in ranks is fairly close to zero. Only about one-fifth of candidates increase their rank by 20 percentile or more, and about half of the candidates do *worse* on their second attempt.

Figure 3.8: Difference in ranks across attempts, Group 4



Notes: This figure uses a sample of candidates who applied for a Group 4 exam for the first time in 2016, were not selected in that exam, and reapplied in the 2017 Group 4 exam. The horizontal axis measures the difference in ranks between the two attempts.

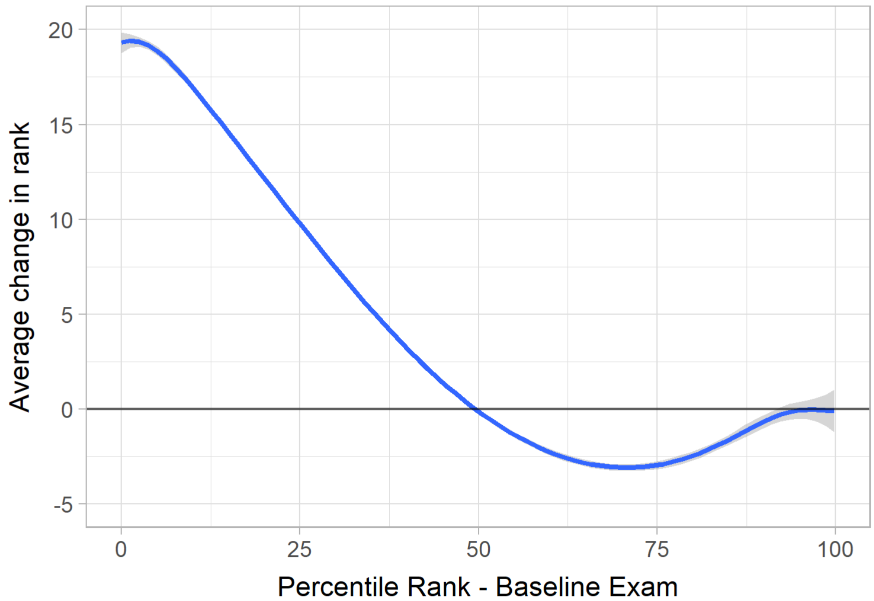
Exam scores exhibit strong mean reversion across attempts

Whether or not candidates are likely to score better or worse in their second attempt depends substantially on how well they perform in their first attempt. As we see in Figure 3.9, candidates with scores below the 50th percentile tend to score better, and candidates with score above the 50th percentile tend to score worse.

The figure tells us that test scores exhibit *mean reversion* across attempts. For most candidates, the main cause of mean reversion is that exam scores reflect a combination of ability and luck.[43] Mean reversion implies that if a candidate got lucky on their first attempt—and end up with a higher-than-usual score as a result—then chances are that they won't get lucky again, and hence their score will be lower on the next attempt. By the same logic, if a candidate got unlucky on the first attempt, chances are they won't get unlucky again, and hence their score will likely increase. This is why we tend to see that scores increase for candidates with low baseline scores (i.e. who got unlucky) and why scores decrease with relatively high baseline scores (i.e. who got lucky).

[43] For candidates at the top and bottom of the distribution, there will also be mean reversion effects due to the fact that percentile rank is a bounded variable.

Figure 3.9: Average change in performance across attempts conditional on baseline performance, Group 4



Notes: This figure uses a sample of candidates who applied for a Group 4 exam for the first time in 2016, were not selected in that exam, and re-applied in the 2017 Group 4 exam. The horizontal axis is the rank in the 2016 exam. The shaded grey area is the 95% confidence interval of the estimate.

The fact that scores exhibit mean reversion throughout most of the distribution means that candidates' true ability remains relatively constant across attempts. However, in the top 30 percentile of the distribution, we can tell the mean reversion effect is weaker, because the curve flattens.[44] This means that among these candidates, there is some genuine growth in ability that counteracts the effects of mean reversion.

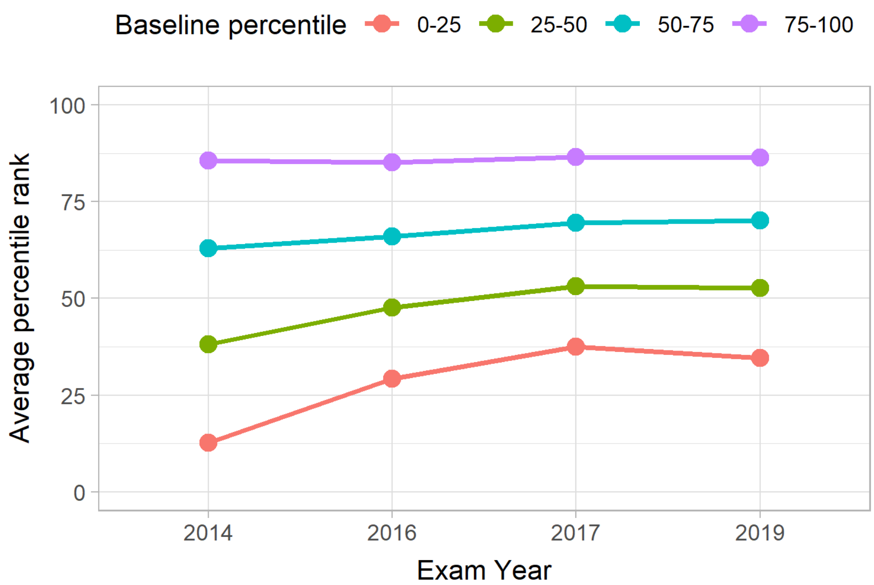
[44] In the case of perfect mean reversion, the line would have a slope of -1, since every point scored above the mean on the baseline attempt would disappear on average in the next attempt. In the bottom 70 percentiles, the slope is -0.37. In the top 30 percentiles, the slope is 0.14.

Candidates' average performance remains roughly constant across multiple attempts

Perhaps there is no average gain from a single additional attempt. Maybe there is a benefit that materializes along longer horizons?

This does not seem to be the case. Figure 3.10 plots the average percentile rank of candidates for a group of candidates who made four consecutive attempts in Group 4 exams, and were not selected in their first three. Candidates are divided into groups based on their performance in their first Group 4 exam.

Figure 3.10: Average rank across attempts, Group 4



Notes: This figure uses a sample of candidates who: i) applied for a Group 4 exam for the first time in the 2014/18 exam; and ii) were not selected in 2014, 2016, or 2017.

There is very little convergence in average performance: candidates who start at the bottom of the distribution remain, on average, at the bottom of the distribution, and candidates who start on the top remain on top.

The main value of additional attempts is that they widen the range of possible outcomes

If most candidates' scores do not improve on average, then where does the value of an additional attempt come from?

The answer is that as candidates make more attempts, the range of possible outcomes expands. Candidates' scores could either increase or decrease by a wide margin. By making additional attempts, candidates are effectively making a bet that this increased variation will work in their favor.[45]

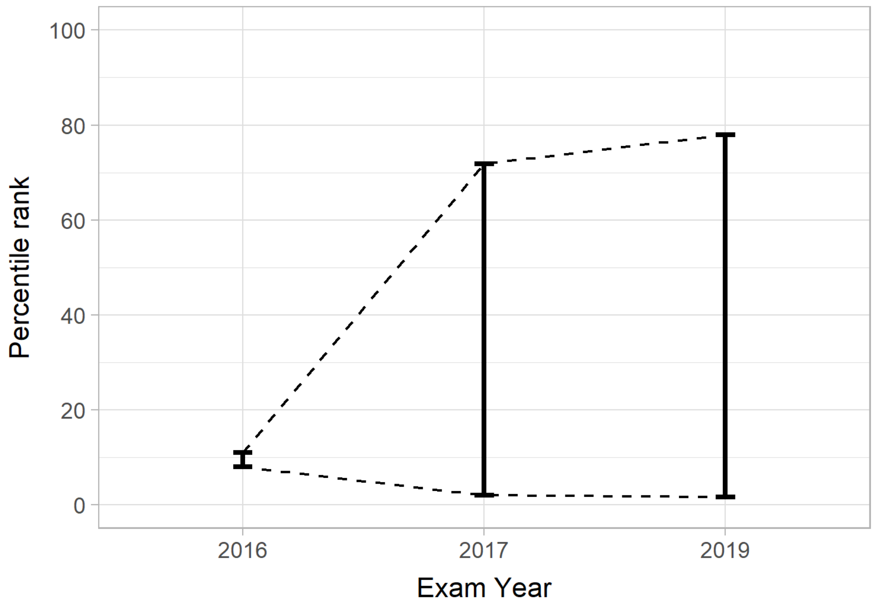
[45] There is an analogy here to securities trading, where people can aim to profit from volatility, even if the underlying value stays the same.

To see how this works, let us consider a set of candidates that scored within a narrow range in their first Group 4 exam, and then track how they fared on their next two attempts (for those who attempted).

The first panel of Figure 3.11 shows that candidates who start at the 10th percentile could end up with even lower performance on their next attempt, or they could even reach the 80th percentile. This extremely wide variation offers hope to those who scored poorly that they may be able to make up for lost ground.

Figure 3.11: Additional attempts widen the range of possible outcomes

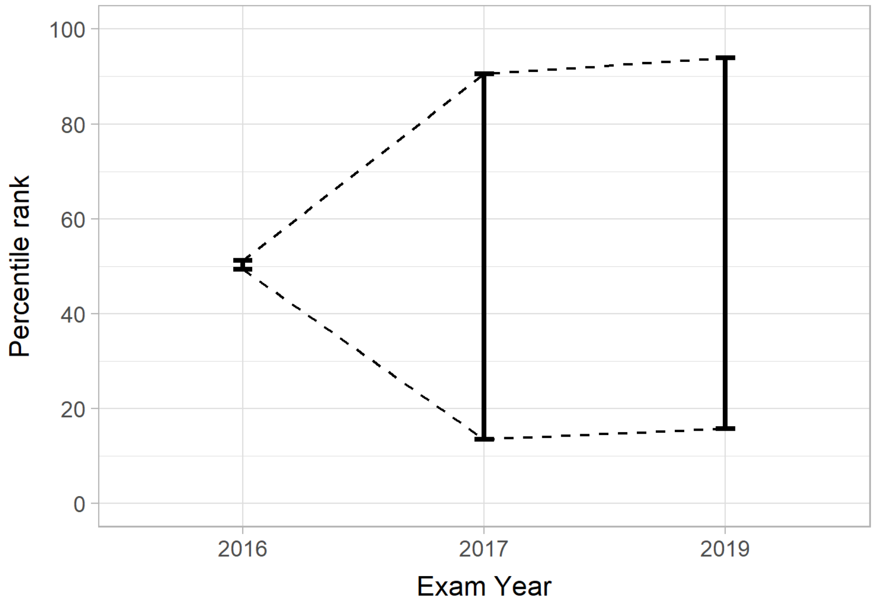
(a) Baseline percentile ≈ 10



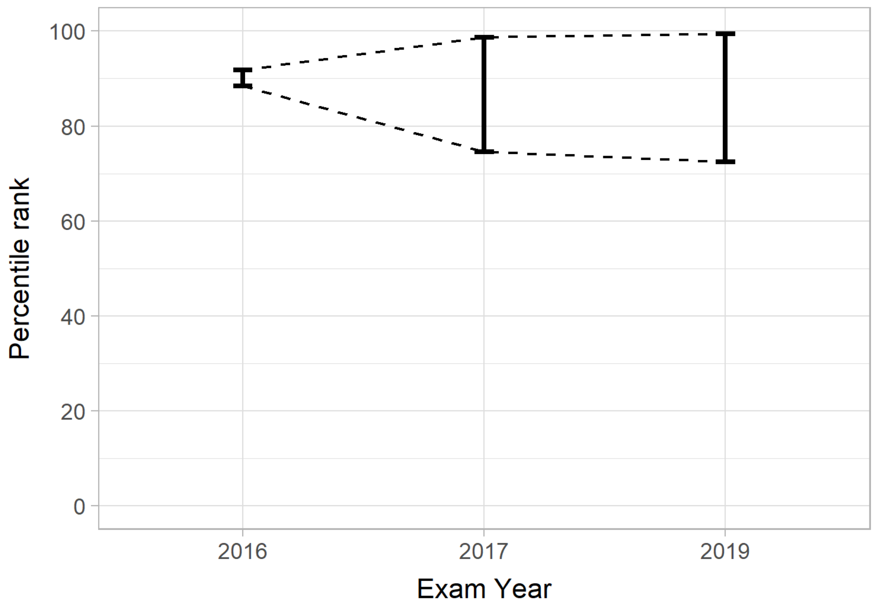
Notes: Figure continues on the next page.

Figure 3.11: Figure 3.11 (cont.): Additional attempts widen the range of possible outcomes

(b) Baseline percentile ≈ 50



(c) Baseline percentile ≈ 90



Notes: This figure uses a sample of candidates who: i) applied for a Group 4 exam for the first time in the 2016 exam; ii) were not selected in that exam; and iii) made at least one subsequent attempt. In each panel, the figure includes candidates whose baseline was within 2 percentile of the given level (i.e. between 8-12, between 48-52, and between 88-92, respectively). The error bars plot the variation in outcomes in subsequent exams for the middle 90 percent of candidates with the given baseline rank.

Meanwhile, for candidates who start at the 90th percentile, the value of future attempts looks quite favorable. There is some downside risk, but there is also a decent chance of making it into the top 1%, which is usually what one needs in order to get selected for a non-typing post.

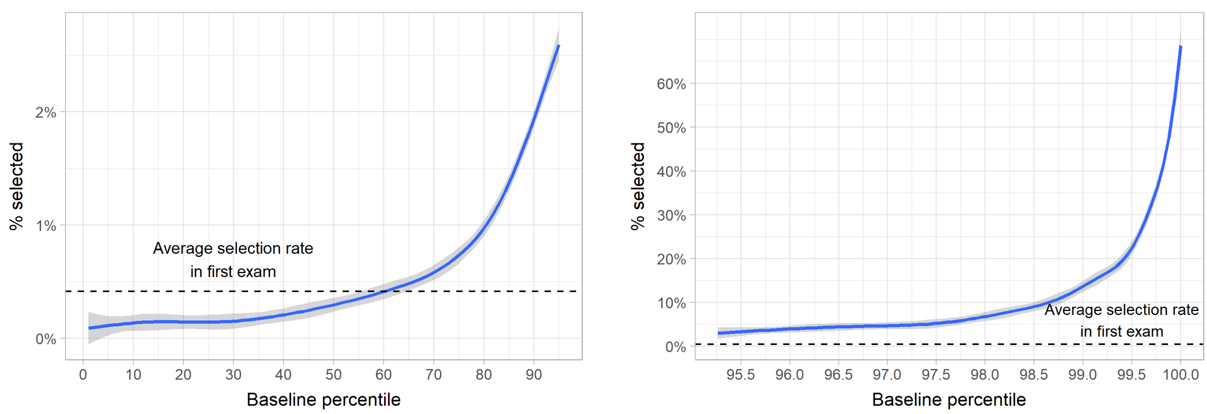
Exam scores on candidates' first attempt are highly predictive of future success

A natural consequence of Figure 3.11 is that most of the candidates who are selected are those who were highly scoring to begin with.

Figure 3.12 looks at a set of candidates who made their first attempt in a 2016 Group 4 exam and were unsuccessful, and looks at what fraction of them were ever selected in any of the subsequent TNPSC Group exams through FY 2019, conditional on re-applying. The figure is split into two because there is so much variation at the top of the distribution.

Figure 3.12: Selection in any subsequent TNPSC Group exam by initial exam score in Group 4

(a) Bottom 95% **(b) Top 5%**



Notes: This figure uses a sample of candidates who: i) applied for a Group 4 exam for the first time in the 2016 exam; and ii) were not selected in that exam. The vertical axis measures the probability of selection in any subsequent TNPSC exam, including non-Group 4 exams. The grey band indicates the 95% confidence interval of the estimate.

For most candidates, the value of making additional attempts remains quite low: for candidates in the bottom 60 percent of the distribution, the probability of eventual selection—even allowing for multiple subsequent attempts—is lower than the average probability of selection in the first exam they took.

However, this picture changes dramatically for top-scoring candidates. For those who scored in the top 5% of the distribution in their first exam, the probability of getting selected in the next attempt can vary anywhere from 3% to upwards of 50% or higher. For these candidates, the risk of participating in the recruitment process looks quite manageable.

Summing Up

Additional attempts are most valuable for candidates who score well in their first attempt. For everyone else, making additional attempts rarely pays off. The reason this is the case is because on average candidates' performance remains constant across attempts. Some candidates manage to increase their scores dramatically, but they are an exception. Most candidates see no meaningful change in their ability, even after years of continuous attempts.

3.5 Uncertainty and misperceptions

From candidates' point of view, the risk of exam preparation depend on how much they know about themselves and the exam process. Do candidates have sufficient information to determine whether they are on track to getting selected? Or do they remain mostly in the dark?

Do candidates know the cutoff?

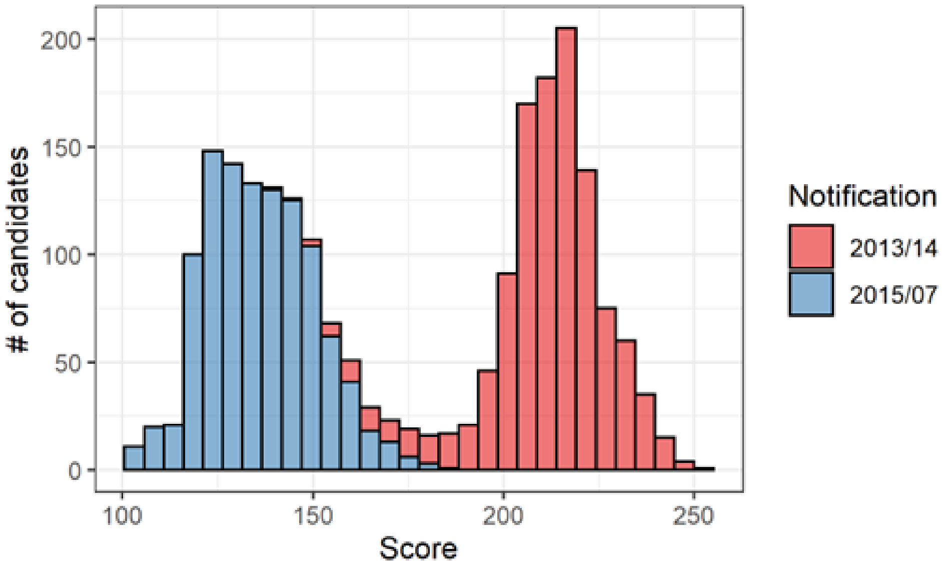
The large variation in cutoff scores from year to year makes it difficult for candidates to predict the cutoff

Exams are normed when their questions are designed so that the difficulty remains constant across different versions of the exam. Norming is a complicated process, requiring extensive specialized knowledge.[46] Because TNPSC, like all other PSCs, does not have a process in place for norming exams, the difficulty of the exams varies widely between iterations, even when they are based on the same syllabus.

[46] See Chapter 5 for a discussion of what it would take for recruitment agencies to start norming their exams.

For example, Figure 3.13 shows the distribution of test scores among successful candidates in two consecutive Group 2 examinations: the 2013 exam, and the 2015 exam.

Figure 3.13: The distribution of test scores among selected candidates in two consecutive Group 2 exams



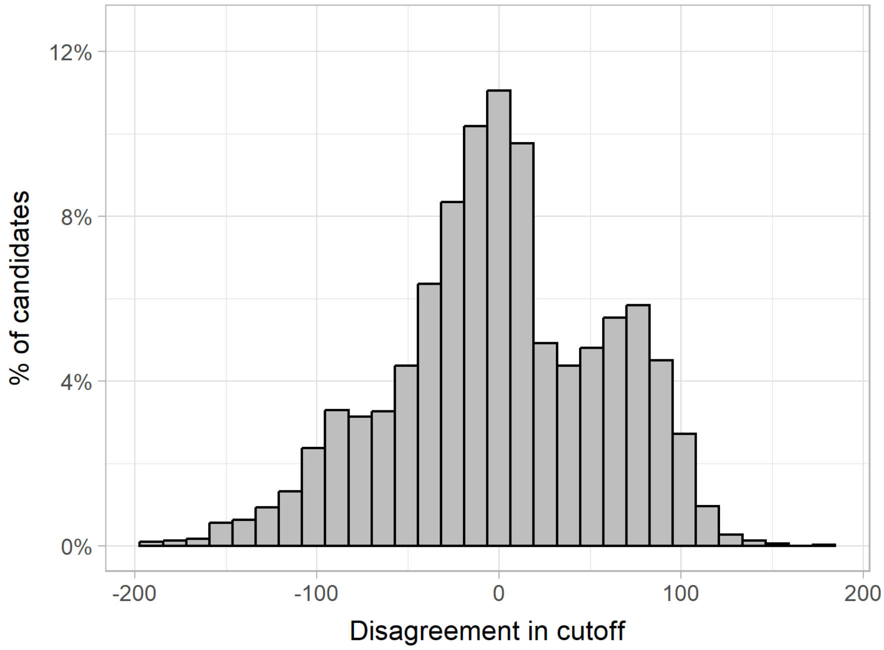
Note that the score distributions are almost entirely non-overlapping. A very good score in one year is a score that is below the cutoff in another year.

It's thus not surprising that candidates are uncertain about the cutoff that will apply to future exams.

In the Candidate Survey, we asked those who were preparing for the upcoming Group 4 exam to predict the cutoff that would apply to their reservation quota (including both community and other reservation criteria). We find that, even among candidates that belong to the same quota, there is substantial disagreement (see Figure 3.14).[47] There are only 300 points on the exam in total, so disagreements of 100 points or more indicate a complete lack of information about what it takes to get selected.

[47] It is also telling that, while piloting this survey question, we discovered that we had to phrase it in a way that explained to respondents what a cutoff was, because some of them were unfamiliar with the concept.

Figure 3.14: Disagreement about the selection cutoff among candidates in the same reservation category



Notes: The figure plots the difference between a candidates' belief on what the cutoff will be and the average belief of all other candidates in the same reservation category, excluding that candidate. Candidates who reported beliefs of either 0 or 300 (the min/max possible score on the exam) are dropped.

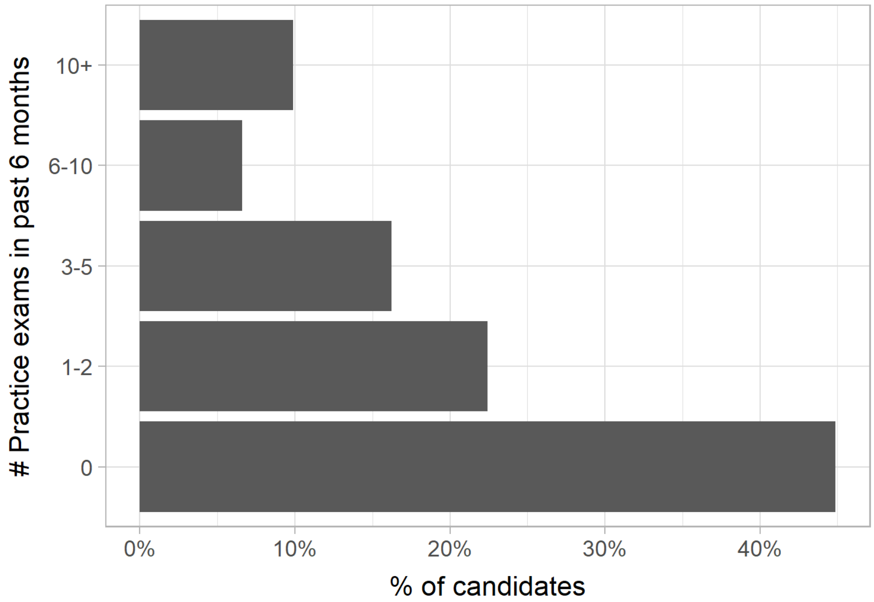
Do candidates use practice tests?

Practice tests may be able to help candidates gauge their progress, but they remain relatively inaccessible.

A well-designed practice test can help candidates determine whether they are on track to get selected.

However, most candidates do not take practice tests regularly (see Figure 3.15). Less than half of candidates had taken a practice test in the 6 months prior to the survey; and an even smaller share have taken them more than once or twice.

Figure 3.15: How many practice tests do candidates take?



Is this because candidates remain unconvinced of the value of practice tests, or because they are inaccessible in some way? Additional research is needed to know for sure; but the fact that practice exam participation is strongly correlated with household income suggests that the cost of practice tests may be a barrier (see Table 3.3).

Table 3.3: Participation in practice exams by household income

HH Income	Took Any Practice Test (%)	Took 6+ Practice Tests (%)
Less than Rs. 5,000	53	17
Between Rs. 5,000 - 10,000	49	13
Between Rs. 10,000 - 20,000	58	16
Between Rs. 20,000 - 30,000	61	17
More than Rs. 30,000	71	36

How much do candidates rely on past test scores?

Candidates mistakenly believe that their past exam performance tells them nothing about their future selection probability

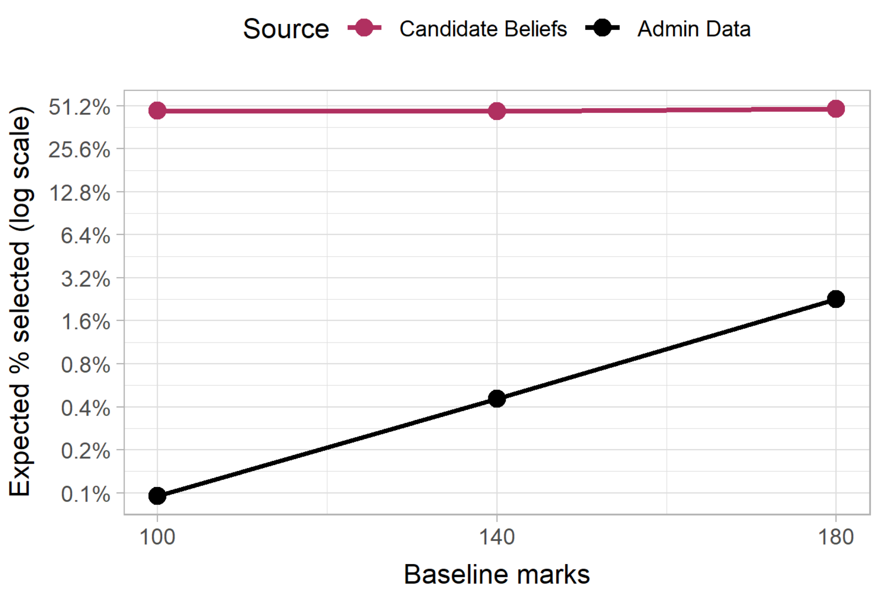
As we saw in Figure 3.12, past exam performance is highly correlated with future success. In the Candidate Survey, we aimed to understand whether candidates also believed the same. Specifically, we asked the following hypothetical question:

> *Suppose in the last Group 4 exam, there are 1000 candidates who get close to [X] marks on their first attempt, but they are not selected for a post. All 1000 candidates re-apply in upcoming Group 4 exam. How many do you think will get selected for a post on this attempt?*

Respondents were randomly shown one of three versions of this question, where in each version the value X was either 100, 140, or 180.[48]

[48] The range of values was chosen so that it would be plausible that candidates would not be selected with those scores, but yet allow for meaningful variation in baseline performance. For reference, relative to the Group 4 exams included in Figure 3.16, these scores correspond on average to the 38th, 86th, and 98th percentiles of the score distribution.

Figure 3.16: Candidate beliefs about future selection rates do not vary with baseline test scores



Notes: Candidate beliefs are based on responses to a hypothetical question in the Candidate Survey. Admin data refers to the average experience for candidates who were making their first attempt in either the 2016 Group 4 exam or the 2017 Group 4 exam, and reapplied in the next Group 4 exam.

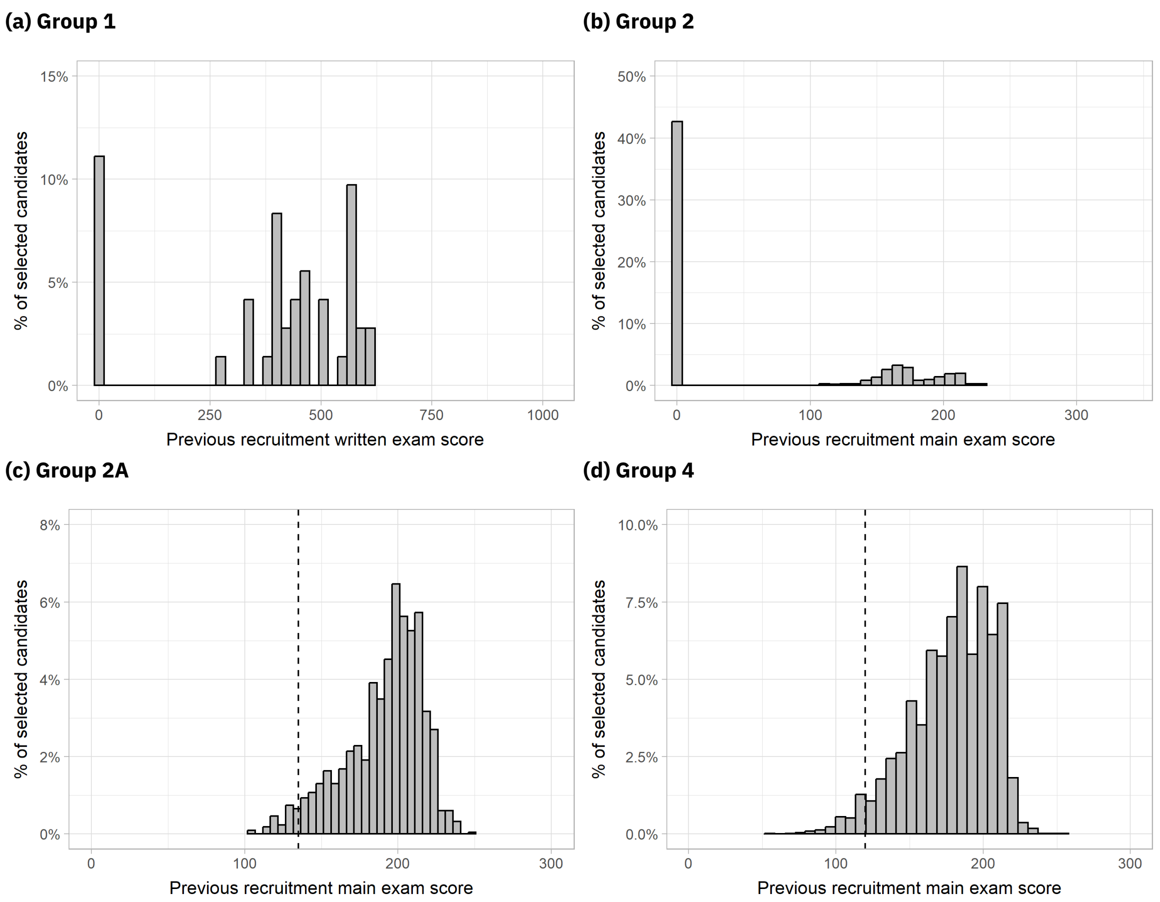
Figure 3.16 plots the results, and compares them to what we observe in the administrative data. There are two striking findings:

- Candidates are way too optimistic about the likelihood of success when re-applying, by a factor of 25 to 500 times.
- Candidates put *no* emphasis on the information available in baseline test scores, even though the observed differential in selection probabilities is close to 25 times.

Success can come from anywhere, even if not everyone is equally likely to be successful.

One of the reasons why it may be hard for candidates to figure out what it takes to be successful is that there are a wide range of profiles that are compatible with exam success.

Figure 3.17: Previous recruitment exam scores for selected candidates



Notes: These figures are based candidates selected in the following exams: Group 1 - 2015/09; Group 2 - 2015/07; Group 2A - 2015/17; Group 4 - 2014/18. For multi-stage exams, the exam score is the sum of the written and interview scores. Any components that are missing due to non-qualification are set to 0.

Figure 3.17 plots the distribution of exam scores for successful candidates in their *prior* attempt. (Individuals who did not appear in the prior exam are dropped from the sample.)

Even among candidates selected in the same exam, their scores on their prior attempt can differ by several hundred points.

However, even though success *can* come from anywhere, that does not mean that it will. As we saw in Figure 3.12, candidates with low scores who are selected are exceptions, rather than the rule. What candidates likely do not realize is that there are many more candidates with low scores who apply compared to candidates with high scores.

3.6 Discussion

How can government policy make it easier for candidates to determine whether they are on track to getting selected?

As it stands, studying for a public sector recruitment exam remains a risky venture, without many opportunities for good, reliable feedback. What policies could help improve this situation?

Providing Detailed Score Reports Recruitment agencies can start to provide candidates with more detailed information about their performance.

For example, in the single-stage exams, TNPSC informs candidates of their total score, their overall rank, and their rank within their communal reservation category.

Here are a couple ways TNPSC can supplement this information:

- They can provide candidates with their *percentile* rank. The overall rank can change dramatically based on the total number of applications, even when candidates' relative performance stays the same. This can make it harder for candidates to keep track of how well they did across attempts. By contrast, the percentile rank is a statistic that is more readily comparable across exams.
- Candidates can be told what the percentile rank is for selected candidates who qualify for the same posts as them, both in the current exam and in previous exams. This can give them a concrete sense of how far away

from the cutoff they are.

- TNPSC has access to information on candidates' performance over time. In principle, TNPSC could include in the score report a table showing how candidates' percentile rank has changed across attempts.
- In addition to providing *individual* information, TNPSC can provide candidates with useful *aggregate* information. For example, candidates should be told just how informative past test scores are for future success. There may be a way of communicating the import of a graph like Figure 3.12 in a simple way.[49]

[49] Of course, providing this information will come as a disappointing shock to many candidates, and that will be challenging to navigate. It will take real skill to figure out how to communicate this well.

These kinds of statistics could help candidates track their progress across attempts—even when cutoffs, vacancy availability, application volume, and exam difficulty keep changing.

Mailing Score Reports to Candidates Even when TNPSC makes information available on its website, it's not necessary that this information reaches candidates. Making information available is the first challenge. The next challenge lies in making sure that candidates pay attention to it.

For example, soon after the examination is over, TNPSC publishes a tentative answer key. This tentative answer key allows candidates to obtain a rough estimate of their final score, which in turn could help them decide whether to continue to prepare.

However, most candidates do not look up their score. In the Follow-up survey, only 30% of respondents who took the Group 4 exam had looked up the answer key and checked their score. This is not just because candidates lost interest in TNPSC, and therefore saw no need in reviewing this information. Even among candidates who said they were planning to apply to a future TNPSC exam, *and were currently studying for it*, the same proportion— 30%—had looked up their score.

The concern is, then, that even if TNPSC were to generate useful information, it may not reach the candidates who need to see it most. One way that recruitment agencies can address this concern is to send score reports by mail to candidates' registered address, in addition to posting them online. This would of course require extra expense and coordination. But it may pay off in other ways. (In Chapter 4, I discuss how recruitment agencies may benefit from additional transparency.)

Ensuring access to quality practice tests Exams happen too infrequently for them to be a useful source of regular feedback. Ideally, candidates would have access to quality practice tests that allow them to gauge their preparedness even before they appear for the exam. This may allow candidates to figure out whether they have a meaningful chance at selection without having to wait for the next recruitment.

The key word here is quality, because practice tests that are misleading (especially those that provide candidates with a false sense of confidence) are perhaps worse than no practice tests at all.

Some recruitment agencies (e.g. Andhra Pradesh PSC) already host online practice question papers on their website. It is worth looking into how these initiatives shape candidate application behavior, and how they can be built upon and improved.

Another important question for future research is how well the existing market for practice question papers functions, and what might be needed to improve it. There are three main concerns to be addressed:

1. Are the practice tests available in the market of good quality (i.e. do they do a reasonable job of helping candidates understand whether they are on track to getting selected)?
2. Can candidates identify the good practice tests, and are those the ones they prefer?[50]
3. Are good practice tests readily accessible?

The fact that public sector recruitment exams are not normed may pose a fundamental challenge to any organization's ability to create a practice test that mimics how a candidate would fare on the real one. In that case, recruitment agencies may have to invest in standardizing their question setting procedures (a topic we take up in detail in Chapter 5) before candidates can have ready access to reliable feedback on their exam performance.

[50] For example, a coaching center director in Maharashtra told us that candidates tend to be attracted to easier practice tests, because it makes them feel good about their exam preparation. These kind of biases in candidate demand could important consequences for candidate application behavior and welfare.

Part II

Using Data to Improve Recruitment Practice

Headquarters of the Tamil Nadu Public Service Commission

4 The Value of Transparency

This chapter is based on joint work with Niharika Singh. What follows is a condensed version of a forthcoming academic working paper.

4.1 Motivation

How transparent Public Service Commissions choose to be in their selection process is an important policy variable

As with any policy decision, recruitment agencies have to balance competing considerations.

On the one hand, the principal risk with increasing transparency is political. In order for recruitment agencies to complete recruitments successfully, the government must be in agreement over how that recruitment should be conducted. In some settings, revealing new information about the selection process can threaten to disrupt the existing agreement, leading to severe delays in recruitment and protracted political fights.

On the other hand, additional transparency could potentially catalyze a number of positive changes in both the recruitment process and in the labor market more broadly:

- For candidates: New information may help them make more informed decisions, and avoid the regret of either investing too little or too much.
- For TNPSC: Proactive disclosure may help build and maintain trust in the selection process.
- For the government and society at large: This information may help reduce the high social cost of the recruitment process (see Chapter 2).[51]

Each recruitment agency weighs the costs and benefits of transparency in its own way. For example, some Public Service Commissions publish all of their annual reports on their website; others do not. Some provide detailed information on selected candidates for current and past recruitments, others do not. Some provide the public with information about candidates that were not selected; most do not.[52]

The problem that recruitment agencies have in deciding how

[51]Outside of public sector recruitments, previous research has shown that providing jobseekers with information can reduce the risk of long-term unemployment (Altmann et al., 2018).

[52]TNPSC has been a leader in proactive disclosure. Under its Open Data Policy (established in April 2022), the Commission is directed to publish *unit-level* data on applicants, test scores, and selected candidates for completed recruitments on its website. All data released under the Open Data Policy is available to download at `https://tnpsc.gov.in/English/OpenDataPolicy.aspx`.

transparent to be is that the risks of transparency are obvious and visible, while the benefits are not as well-documented. Recruitment agencies usually do not have a reliable mechanism for receiving feedback from candidates, especially for policies that are expected to produce diffuse and incremental benefits. As a result, we just don't know how much these additional transparency measures could help.

How do candidates respond to increased transparency?

To answer this question, we ran a small pilot experiment. We provided a sample of potential candidates with new information that was not otherwise publicly available. This was meant to simulate what it might be like for a recruitment agency to initiate a new transparency measure from candidates' point of view. We then compare the beliefs, attitudes, and behavior of candidates who received this new information to another sample of candidates who did not.

The experiment was conducted as part of the Candidate Survey. Our sample consists of 949 candidates who were preparing for the upcoming Group 4 exam, and who had made less than two prior attempts.[53] We randomly assigned 464 candidates to the Treatment group, and the remainder to a Control group.

Our main outcomes of interest are based on data that was collected 5-7 months after we showed them the information, and 3-6 months after the Group 4 exam was over.

4.2 The Information Experiment

Experimental Design Since the survey was conducted online, we relied on the survey platform to automatically randomize respondents into Treatment and Control Groups. In Appendix B.3, we confirm that Treatment and Control groups are balanced on key characteristics.

Candidates were shown information about the previous Group 4 recruitment (i.e. the one notified in FY 2019). In the Control group, respondents were shown information about: i) the total number of candidates appearing; and ii) the total number of candidates that were "newly selected."[54] This is mostly public information: it is already published in newspapers and can be found in the Annual Reports posted on TNPSC's website.[55]

In the Treatment group, respondents were provided with the

[53] This sample is part of a larger study with 3,574 individuals. We focus our attention on this sub-sample in order to simplify the analysis. For more details on how the analysis sample was constructed, see Appendix B.3.

[54] By "newly selected," we meant that we were providing information on candidates who did not already have a government job. Our intention was to communicate to respondents (who we expected not to have a government job) that the information we were presenting was directly relevant to them.

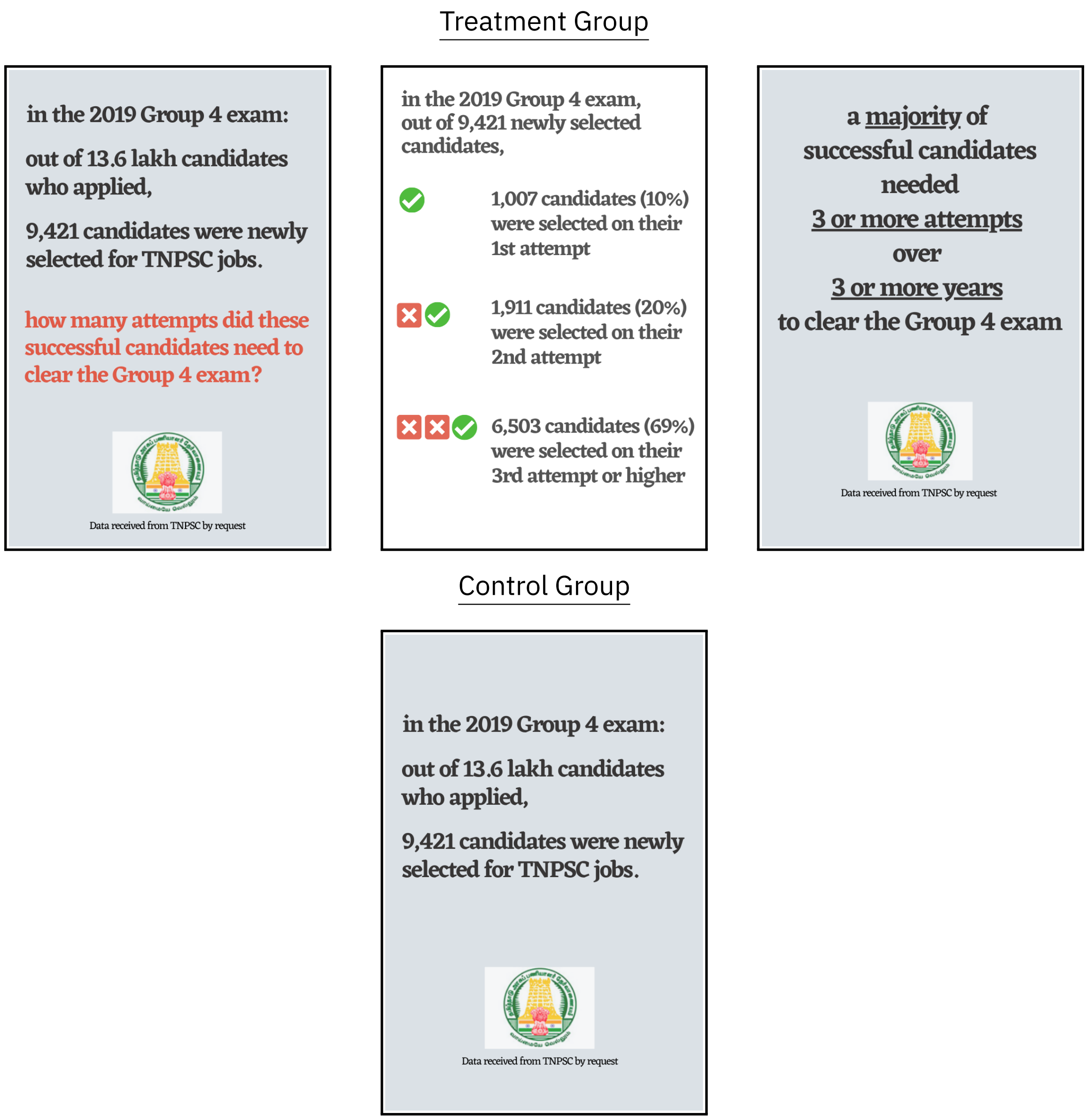
[55] Technically speaking, public sources do not separate statistics by whether candidates already have a government job. But the statistics we present are nearly identical to the publicly available ones.

same information as in the Control group, plus some new statistics that we calculated using the administrative data. Specifically, candidates were told the distribution of the number of prior attempts that "newly selected" candidates in that recruitment had made.

How was the information presented? Based on their treatment status, candidates were shown a different set of information cards while they were filling out the survey (see Figure 4.1). The information was provided in either English or Tamil, depending on the language that the respondent chose to complete the survey in. The cards aimed to communicate the information in visual, intuitive manner.[56]

[56] To ensure that candidates paid attention to the information, we included several "attention checks" in the survey. Right after the candidates were shown the information, they were asked quiz questions. If candidates responded incorrectly, they were reminded of the correct answer.

Figure 4.1: How information was communicated in the Information Experiment



How did we decide what information to show?

There are two components to candidates' beliefs about their selection probability:

- *Extensive margin beliefs* are beliefs about the likelihood of getting selected in the first place, i.e. "what is the chance of ever getting a government job?"
- *Intensive margin beliefs* are beliefs about how much investment it will take to get selected, i.e. "how long will it take to get selected?"

While conducting background research for this experiment, we found that candidates tended to react very differently to the information we presented, depending on whether we targeted extensive or intensive margin beliefs.

In general, candidates reacted negatively to information about the extensive margin. In interviews and focus groups, it was not uncommon for candidates to get angry with us for bringing up the possibility of not getting selected. Some candidates responded by disengaging from the conversation. There was a tendency to refuse to contemplate the possibility of not getting selected. If forced to discuss, candidates could easily dismiss the statistics we showed them by making some combination of the following arguments: i) most candidates were not studying seriously, but they were; ii) even among candidates who were studying, many were not following the right method of studying, but they were; iii) if you think negatively, then the selection definitely will not happen, so you have to think positively. As we saw in Chapter 3, these rebuttals are not ill-informed.

It's important to understand this reaction in context. Most candidates are constantly discouraged by the people around them. Even if parents are supportive, other neighbors and relatives will taunt them, call them good-for-nothing, find ways of comparing them with others, or shame them for their unemployment. It can be a constant struggle to stay motivated. When we, as relative strangers, talk to candidates about the possibility of not getting selected, we are just another one of those negative voices—someone who is seen as not having their best interests at heart. It is natural that candidates would have a defensive mechanism to resist this negativity, otherwise it would be difficulty to put in the effort needed to ever succeed in the first place.

Our ability as researchers to communicate the possibility of not getting selected is thus quite limited. We therefore stuck to the more palatable discussion topic of how long it will take to get selected. Candidates were more willing to engage with this topic, because it was less threatening. But even though we did not bring up the possibility of not getting selected directly, this topic of conversation still manages to allude to it. What was implicitly understood is that if it takes too long to get selected, candidates may be forced to drop out before they get selected.

4.3 How accurate are candidates' beliefs?

Most candidates underestimate the number of attempts that successful candidates need to make.

As part of the Candidate Survey, we asked respondents to make a guess about how many previous attempts successful candidates made in the 2019 Group 4 exam. This question was asked before candidates were shown any information.

The exact wording of the question was:

> *There were 9,421 people who were newly selected for a TNPSC post in the 2019 Group 4 exam.*
>
> *Out of these 9,421 selected candidates, how many do you think were selected only after 3 or more attempts on a Group 4 exam?*

Candidates responded by dragging a slider on the online application form somewhere between 0 and 9,421. The response was converted into percentage terms by dividing by 9,421. We then compare their answer to the correct answer. (The correct answer is 69%.)

Figure 4.2 plots the discrepancy between candidates' beliefs and the correct answer.

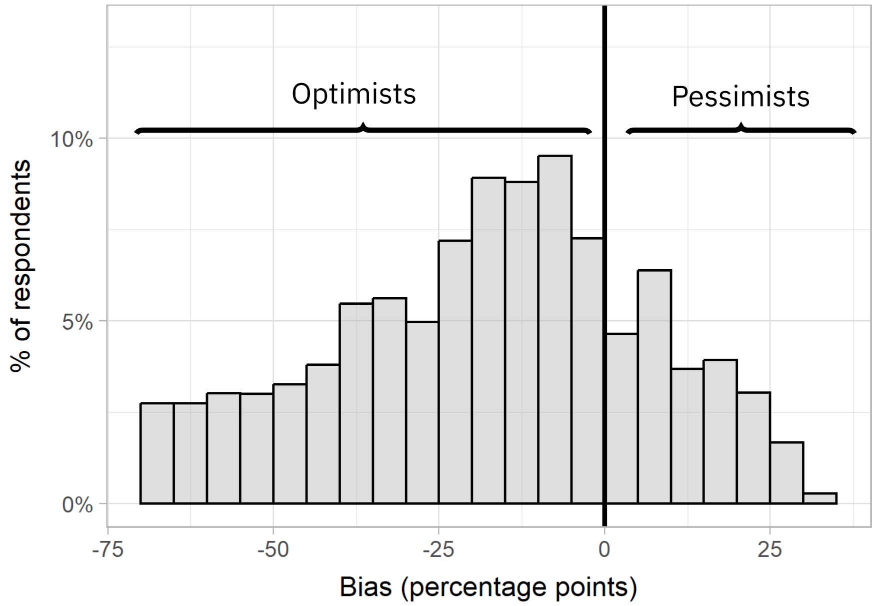
We'll call candidates *Optimists* if they tend to believe that successful candidates need fewer attempts than they really do (and hence they're more optimistic of getting selected themselves in under three attempts). We'll call the candidates *Pessimists* if they tend to believe that successful candidates need more attempts to get selected than they really do.

Most candidates are too optimistic, and by a wide margin:[57]

- Overall, about *three-fourths* of candidates are Optimists.
- The average bias is -16 percentage points, i.e. on average candidates believe that 53% of candidates are selected in 3 or more attempts, when the truth is 16 percent points higher.
- Only 26% of candidates are able to guess within 10 percentage points of the correct value.

[57] Over-optimism in job search is not limited to public sector recruitment, but appears to be a general feature of unemployment in the labor market (Mueller et al., 2021).

Figure 4.2: How biased are candidates beliefs about the share of candidates that are selected in 3+ attempts in the 2019 Group recruitment?



Notes: About 6% of respondents report prior beliefs of either 0% or 100%. This is likely due to a misunderstanding of the question. These responses have been dropped from this figure.

Table 4.1 compares Optimists and Pessimists on a range of observable characteristics. Both groups have similar age and gender profiles. However, Optimists are less likely to have prior exam experience, and are more likely to have done well in school exams.[58]

[58] This pattern raises the possibility that, in the absence of experience, candidates extrapolate the difficulty of TNPSC exams from how difficult they found school.

Table 4.1: How do Optimists and Pessimists differ?

Variable	Optimists	Pessimists
Panel A: Demographics		
Average Age	24.4	24.9
Female	49%	44%
Panel B: Education		
College Graduate	52%	54%
Average 10th Std Board Marks	374	360

Notes: Table continued on the next page.

Table 4.1 (cont.): How do Optimists and Pessimists differ?

Variable	Optimists	Pessimists
Panel C: Exam Experience		
Average # Previous Group 4 Attempts	0.16	0.27
Took any previous TNPSC Group Exam	44%	56%
Panel D: Time Use Before the Exam		
Working	39%	46%
Average Study Hours per Week	12.64	12.91

Notes: Each column tabulates the average value of the variable for Optimists and Pessimists, respectively.

Candidates who are confident in their beliefs are not necessarily more accurate

Do candidates know how much they do not know? To answer this question, we asked respondents to tell us how confident they felt about their guess, on a 5-point scale ranging from "Not confident" to "Very confident."

Unfortunately, candidates appear to not recognize how uninformed they are. About 45% felt "Confident" or "Very confident" about their beliefs. There is very little correlation between a candidates confidence, and the accuracy of their beliefs. And even the most confident candidates had beliefs that were off by 22 percentage points on average.

4.4 Impacts on candidates' beliefs

Candidates use information from prior exams to make inferences about the difficulty of future exams

Right after we showed respondents the information cards, we asked them to predict the share of successful candidates in the upcoming exam who will have made three or more attempts. Specifically, we asked:

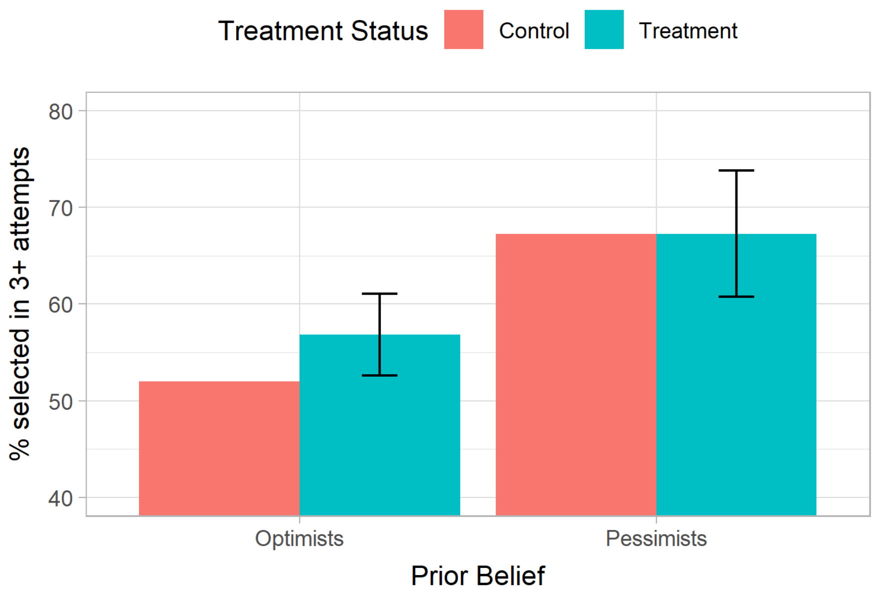
> *In the upcoming Group 4 exam, about 7,500 candidates will be selected. Out of these 7,500 selected candidates, how many do you think will get selected only after 3 or more attempts in a Group 4 exam?*

Figure 4.3 summarizes the impact of treatment. We show impacts separately for Optimists and Pessimists to account for the fact that, in theory, these two groups should revise their beliefs in opposite directions based on the information.

Beliefs are responsive to the information treatment. This tells us that candidates do in fact use and incorporate information from prior recruitments when forming beliefs about future recruitments.

The effect of treatment is concentrated on Optimists. In response to treatment, these candidates are more likely to expect the exam will favor candidates with prior experience. Pessimists' beliefs remain roughly in place. This might be because even without the information, candidates' beliefs are roughly in line with what was shown on the information card, so providing information may have simply reinforced the existing belief.

Figure 4.3: The effect of information on candidate beliefs before the exam



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

The experience of taking the exam provides a large information shock of its own

After taking the 2022 Group 4 exam, candidates now have two sources of information about how difficult it might be to be selected: 1) beliefs based on prior exams; or 2) experience based on the exam they just took. Is the prior information still relevant? How do these different sources of information interact with each other?

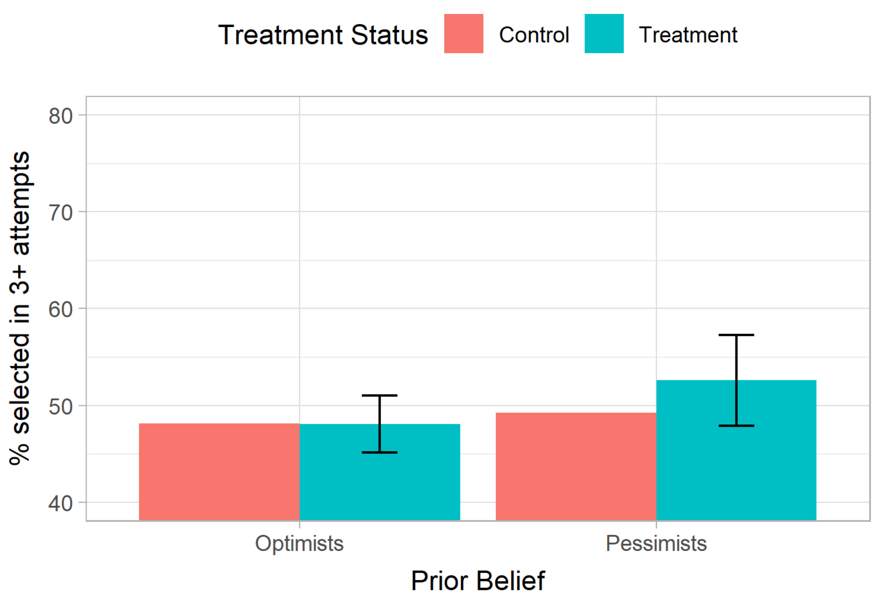
We address this question by measuring beliefs again in the Follow-up Survey. We asked respondents a question analogous to what we asked in the baseline survey:

> *There are roughly 7,500 candidates who will be selected in this Group 4 exam. How many of them do you think were selected after 3 or more attempts?*

As before, we estimate the fraction of candidates that they believe are selected in three or more attempts by dividing their response by 7,500.

Figure 4.4 plots the impact of treatment on candidates' beliefs *after* they have taken the exam.

Figure 4.4: The effect of information on candidate beliefs after the exam



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

Note that the level is substantially lower. This means that candidates generally found the exam to be much easier than they expected (and hence fewer candidates would need 3+ attempts). It is likely that the 2022 Group 4 exam was particularly informative for candidates because the syllabus underwent a major revision for the first time in nearly a decade.

In light of the new information that the experience of taking the exam provides, the treatment effect largely vanishes: after the exam, candidates in the Treatment and Control groups hold similar beliefs.

4.5 Impacts on candidate behavior

Even though the information cards may not have left a lasting effect on the beliefs that we measure, they may still have had an effect on candidate behavior. For instance, the information could have shifted candidates' beliefs about future exams, or it could have shifted the beliefs of other decision-makers in the household.

There are several different dimensions along which candidate behavior might change:

- **Study Behavior**: Whether candidates continue to invest in exam preparation *after* the exam is over
- **Intentions**: The number of attempts that candidates plan to make in the future

---

Study Behavior

The information cards increased labor force participation and employment rates

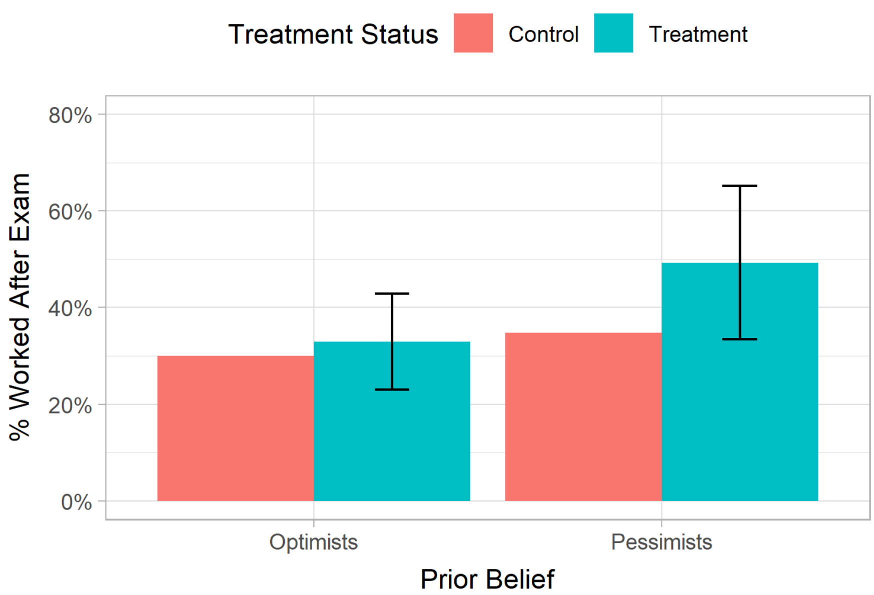
A perhaps surprising result is that the information appears to have encouraged some candidates to start searching for jobs and find employment.

The effect is concentrated on Pessismists, i.e. those who expected it to be difficult to get selected. Figure 4.5 shows that among Pessimists there was about a 38% increase in the share of candidates who had searched for a job after the Group 4 exam concluded. Figure 4.6 shows that among Pessimists there was a 42% effect on the share of candidates who had worked in any job since the exam concluded. Averaged across the whole sample, this corresponds to effects of 11% and 12%, respectively.

Figure 4.5: The effect of information on the share of candidates who have searched for a job since the Group 4 exam concluded

Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

Figure 4.6: The effect of information on the share of candidates who have worked since the Group 4 exam concluded



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

Intentions

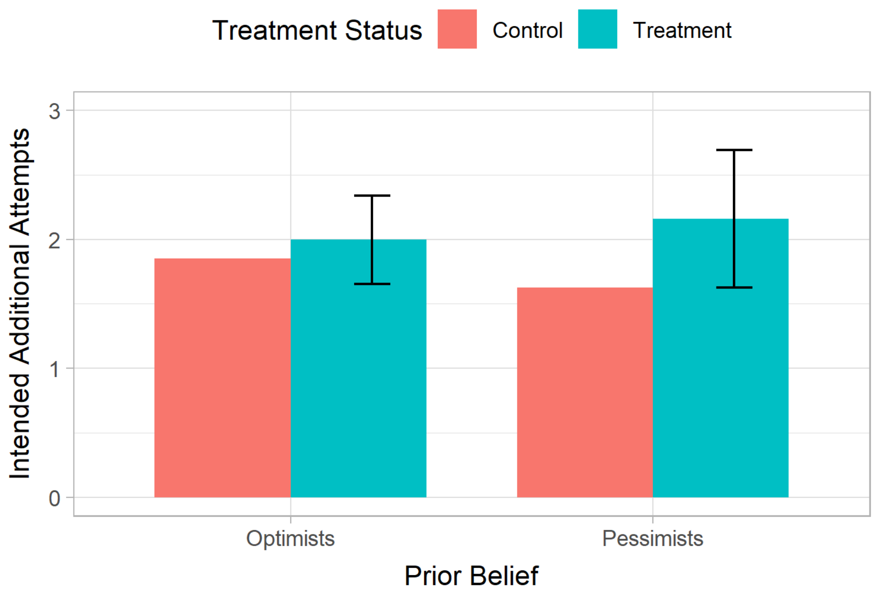
On net, the information encourages candidates to make additional attempts

It is too soon to determine whether the information affects the number of attempts they will make. (For that we would have to wait until the next TNPSC exams are held.) But we can look at what candidates intend to do. In the Follow-up Survey, we asked candidates:

> *If you are not successful, how many more attempts do you think you will be able to make for the TNPSC Group 4 exam?*

Figure 4.7 plots the effect of treatment on candidates' responses. For Optimists we see no meaningful change in intentions. For Pessimists, there is a 33% increase in the average number of desired attempts, from 1.63 to 2.16. Because Pessimists in the minority, the overall effect is more muted. On net, the number of future attempts per candidate is expected to increase by 9%.

Figure 4.7: Effect of information on number of additional attempts candidates intend to make



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

4.6 Impacts on candidates' attitudes towards TNPSC

There is suggestive evidence that providing additional information improves candidates' approval of TNPSC

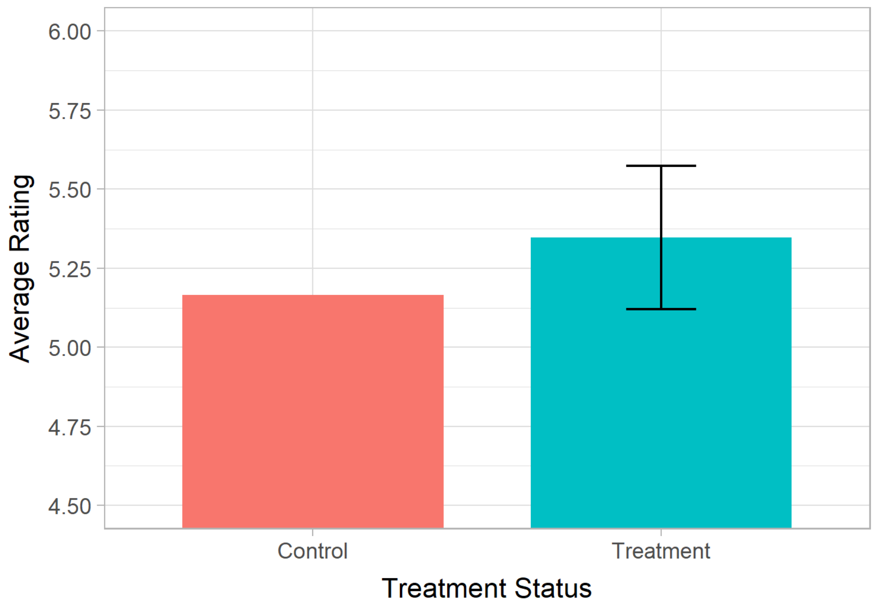
In the Follow-Up survey, we asked the following question:

> *On a scale of 0 to 10, how would you rate TNPSC's management of the exam process? (0 means very bad, 10 is excellent)*

Figure 4.8 shows the effect of treatment on how candidates answered this question. Here, we combine the effects of Optimists and Pessimists, since they move in similar directions.

We find some suggestive evidence that TNPSC's rating *increased* by about 0.13 standard deviations as a result of the treatment.

Figure 4.8: The effect of information on candidates' rating of TNSPC management



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

This suggests that TNPSC is unlikely to suffer from providing useful information for candidates—and may even be able to enhance its reputation by doing so.

4.7 Do candidates value the information we provided?

73% of candidates are willing to give up a Rs. 50 gift card for the information that we provide.

Respondents react to information that is provided. But does that mean they value it? Do candidates knowingly treat this kind of information as relevant to achieving their goals?

One way to find out is to see whether they are willing to give up something else of value for it.[59] After respondents completed the survey, they were asked to choose a gift item. The gift item was meant to thank them for the time and effort it took to complete the survey. The two choices were:

- **Rs. 50 Amazon Gift Card**: This is meant to be a near-substitute for cash. The gift card can be can be added to an Amazon wallet and used for most types of online purchases
- **A soft copy of the information card**: We offered to email / Whatsapp a PDF copy of whatever information card that candidates saw as part of completing the survey.

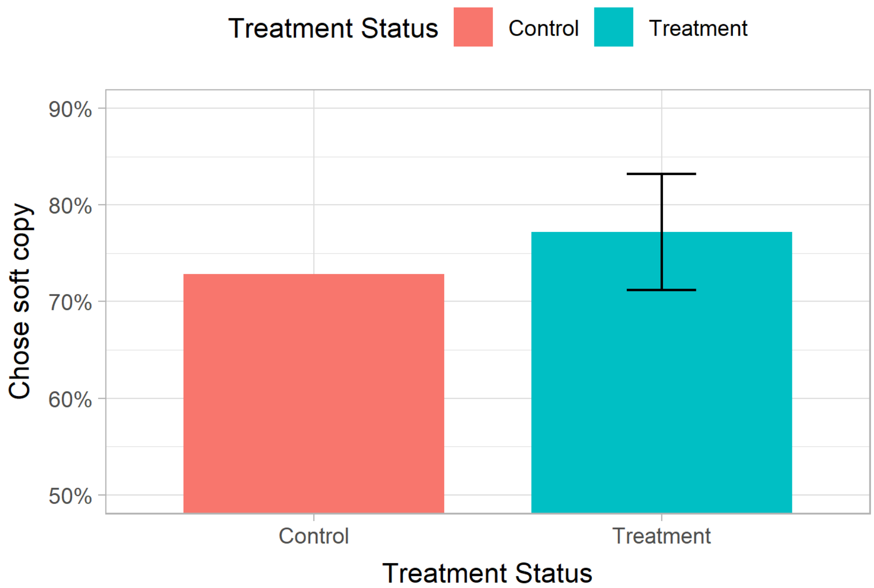
Someone who chooses the soft copy of the information card values it at least as much as Rs. 50. For this sample, this is a significant sum: on average, candidates in the sample were earning Rs. 151 per day.

How often do candidates prefer the soft copy of the information card? Figure 4.9 graphs the take-up rate by treatment status. The high-take up rates of over 70% suggests that there is remarkable unmet demand for information in the candidate pool.

Curiously, we see practically no difference in the demand for information cards between Treatment and Control groups. In other words, just telling candidates aggregate statistics seems to be just as valuable as more detailed information on who gets selected.

[59] Asking respondents directly whether they value the information may not produce reliable data. This is because of a phenomenon known as "social desirability bias," wherein a survey respondent feels a need to be polite to the researchers (i.e. in this case thanking them for providing the information) rather than reporting what they actually believe.

Figure 4.9: Take-up rates of the soft copy of the information card by treatment status



Notes: The treatment effect is estimated using the regression specification detailed in Appendix B.3. The error bars indicate 95% confidence intervals, a measure of the uncertainty of the estimate.

Who values the information the most? Table 4.2 compares the characteristics of candidates who preferred the information card over those who preferred the gift.

Table 4.2: Who is willing to give up Rs. 50 for the information card?

Variable	Chose Information Card	Chose Gift Card
Panel A: Own Access to Information		
Spends on Coaching Classes	12%	12%
Has No Personal Connections to Selected TNPSC Candidates	67%	67%
Panel B: Household Access to Information		
HH Income Less than 10K per month	57%	44%
Mother Not Completed 10th Std	76%	62%
Father Not Completed 10th Std	71%	57%

We find very little correlation between measures of candidates' own access to information and take-up of the information cards. Instead, it seems that the candidates who chose the information card are more likely to come from disadvantaged households where *other* members' access to information may be less. This raises the possibility that candidates preferred to have the information card because they wanted to show it to other household members.

4.8 Discussion

Transparency in public sector recruitment can be an instrument for active labor market policy

The findings of this pilot study suggest that employment rates are constrained not just by demand but also by supply. This provides an opening for information provision to have real effects on labor market outcomes.

The size of the effect can be quite large, especially when multiplied across the whole population of candidates. Even if the effect of a scaled-up policy were half as large, providing candidates with information about selection rates could potentially increase the number of employed people in the labor market by 1 lakh. For reference, this is roughly comparable to the total number of people in Tamil Nadu placed through vocational training programs through the government's flagship skill development initiative between 2016 and 2020 (MSDE Annual Report 2021-2022).[60]

[60] There were 1,23,476 people in Tamil Nadu placed through either Short Term Training (STT) or Special Projects (SP) under PMKVY 2.0, through both the Central and State components of the scheme.

Providing information to candidates should be complementary to other efforts to improve conditions in the labor market

At best, information can only help candidates better understand the risks of their decisions. On its own, it cannot improve the opportunities available to candidates, nor does it address the underlying yearning that candidates have to improve their economic well-being. Providing information can help direct candidates towards more profitable investments; but other policies need to be in place first to ensure that those alternatives exist.

However, there are ways in which information interventions can *complement* other reforms aimed at improving the quality of jobs available. For example, one of the primary sources of good employment around the world is jobs that use specialized, technical skills. Over the past decade, India has made a substantial push to provide young people with these skills, but these programs have suffered, in part, from a lack of interest. A report commissioned by the Tamil Nadu Skill Development Corporation notes that one of the challenges of attracting candidates for vocational education is that many of them are either waiting on the results of a competitive exam they took, or are preparing for a future competitive exam (Tamil Nadu Skill Development Corporation, 2019).

Done well, a targeted information intervention can help speed up the adoption of high-quality, viable vocational training programs. This kind of information may be especially valuable for candidates who themselves are keen on trying some of the newly available careers in the market but need help convincing their more traditional family members to take this less familiar path.

However, targeted information interventions can be misused as well. Some candidates can be quite impressionable. Knowing this, unscrupulous actors can take advantage. For example, telling young people they will never get selected in a government job could be a way of scaring them into signing up for low-quality vocational training programs as well.

The details of how information is provided therefore matter considerably. The underlying reason why these details matter is that they have distributional consequences. When candidates change their behavior because of information, they free up time and money that would have otherwise gone into exam preparation for other potential uses. This leads to a question of who captures that value. If information drives candidates to give up, to feel hopeless about finding better opportunities, to take whatever job they can find—that value can be captured by people who want easier access to candidates' time and money. If, on the other hand, the information helps empower candidates to find other, more reliable ways of investing in their livelihoods, then candidates can capture some of that newly created value as well.

What is the best way of providing information to candidates at scale?

Issuing press releases alone probably will not be sufficient. It takes a well-functioning communication ecosystem to make sure that accurate information reaches candidates in all corners of the state, and is properly understood.

At present, the work of informing candidates is largely shouldered by coaching institutes. Institutes provide substantial amounts of free content online—especially through Youtube, Facebook, Telegram, Instagram, and other forms of social media—explaining not just study strategies, but also important logistical information, such as how the cutoff is determined, or how to fill out the form online.

However, it is unlikely we will be able to rely on coaching centers to provide candidates with accurate information about selection probabilities. (Instead, their content tends to feature candidates who have passed in a single attempt, and offer students tips for how they can achieve the same).

In that case, there may be a case for recruitment agencies, and the government more broadly, to step in to provide this information. A concerted effort by the government, academics, and organizations representing candidates' interests could find ways of developing solutions to the problem.

5 Improving Exam Quality

5.1 Motivation

Exams vary in their quality.

An exam is a measurement tool, like a weighing scale or a ruler. Exams measure some latent trait in individuals, which is often referred to as ability or merit.

The following is a list of some of the many dimensions of quality that matter in a public sector recruitment exam:

- **Predictiveness**. We would like to make sure that the trait that the exam is measuring also predicts how well that candidate will perform as a government employee.
- **Differentiation**. We want to make sure that the exam generates a wide distribution of scores, so that we can more easily distinguish candidates with high ability from those with less ability.
- **Consistency**. Someone with the same ability should get similar scores on different versions of the same test. This way they can easily track their progress over time, and focus their studies in the correct direction.

Some dimensions of exam quality can be measured with the data that recruitment agencies already have.

In this chapter, we explore how insights from a paradigm within psychometrics known as *Item Response Theory* can be used to quantify exam quality. This is the same approach that is currently used in well-known assessments such as the *GRE* and the *TOEFL*.

The advantage of quantifying question quality is that it allows recruitment agencies to track their progress and make systematic improvements, rather than relying on the subjective opinion of Commission members to drive change.

Not all dimensions of quality are amenable to quantification. For example, predictiveness is quite important, but there are no existing measures of performance or productivity for most government employees, let alone a measure that could tell us the potential performance of not-selected candidates.[61]

However, recruitment agencies can make substantial progress on measuring differentiation, using the data that is already

[61] Some studies show that *among candidates who are selected*, those who score relatively better are more productive (Dahis et al., 2020; Xu et al., 2021). But this does not answer the question of whether the exam is selecting on the right traits to begin with.

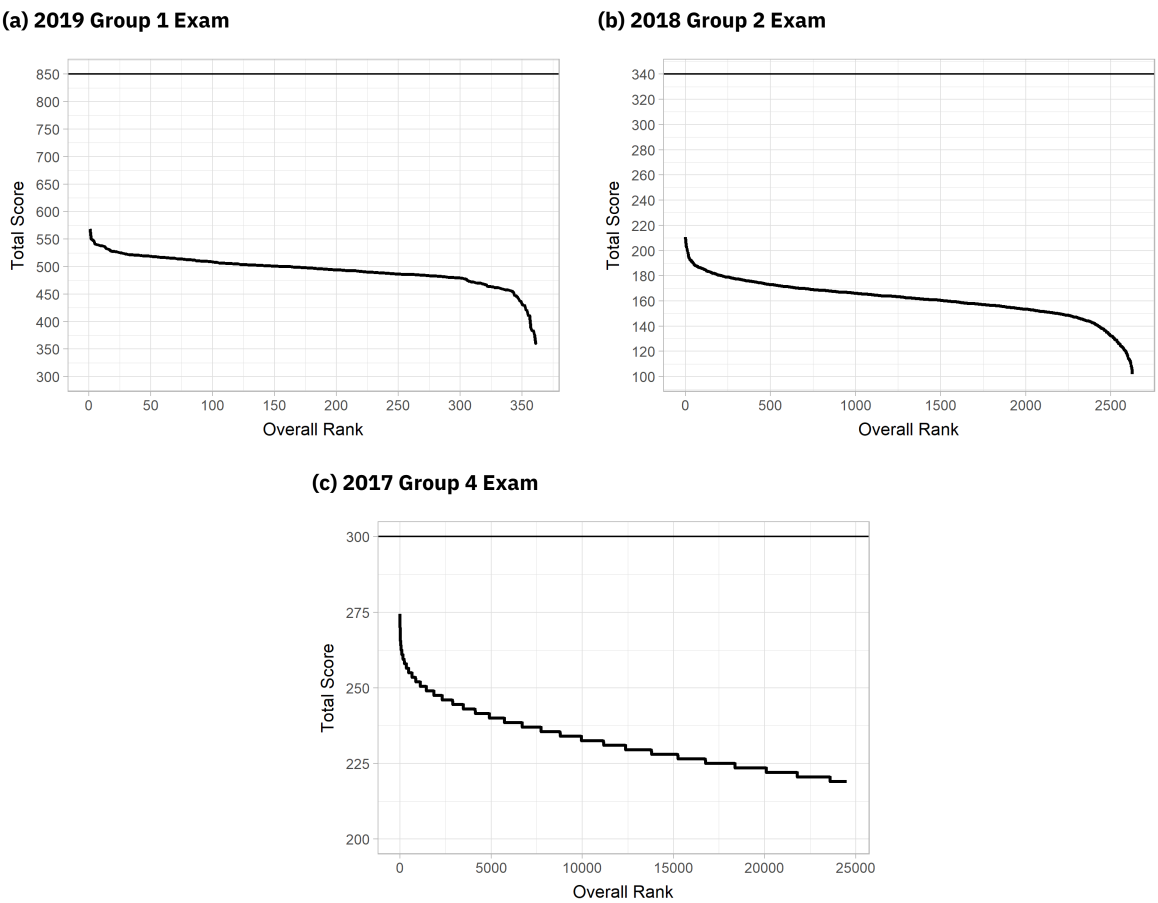
generated by the testing process. This is still useful. How well a selection exam differentiates between candidates is closely related to the purpose for which the exam was created. Given how high application rates are, selection exams are expected to distinguish ability among lakhs of applicants. To meet this requirement, the difference in test scores between candidates of different ability levels should be wide, especially at the top of the distribution. It's therefore important to understand how well the existing framework for setting exams is able to measure up to this standard, and identify where improvements can be made.

The distribution of exam scores is highly compressed, especially at the top of the distribution

Figure 5.1 illustrates how compressed the score distributions currently are using recent exams as examples.

Figure 5.1: Score compression at the top of the distribution

(a) 2019 Group 1 Exam **(b) 2018 Group 2 Exam** **(c) 2017 Group 4 Exam**



Notes: For Groups 1 and 2, the figure plots total marks for all candidates who appeared for the oral test. For Group 4, the figure includes candidates with rank up to twice the number of available vacancies. The black line on the vertical axis marks the maximum possible score.

There are two features of these graphs that draw attention:

- In both the Group 1 and Group 2 exams, the test score distribution occupies a small fraction of the available variation. The highest score that any candidate obtains is about 75-100% lower than the maximum possible score. This means that there is a lot of available room to stretch the score distribution further. (For Group 4, the top test scores are already close to the maximum possible score.)
- The range of scores among top-scoring candidates is not wide. For example, in the Group 4 exam, there are only 51 marks that separate the candidate with rank 1 from the candidate with rank 20,0000. As a result, on average each candidate within this range shares the same score as 1094 other candidates.

These patterns have remained consistent across the exams conducted for Group Recruitments between FY 2013 and FY 2019.

What is the value of improving quality? What would it take to do so?

Improving question quality in a systematic and long-lasting way will likely be an ambitious project. Is it worth the effort? This question deserves serious consideration. This chapter discusses some of the major costs and benefits of improving exam quality, as well as the practical issues that would arise if recruitment agencies were to take on this challenge.

5.2 The value of exam quality

Why does score compression at the top of the score distribution matter?

When the score distribution is compressed, the exam process is likely to be more fragile.

Because of the compression of the distribution, a difference of a single point makes a big difference in selection outcomes. This creates an incentive for unsuccessful candidates to dispute the selection results in the courts. As a result, Public Service Commissions are typically inundated with court cases during and after the recruitment process.

Lawsuits can be a serious threat to a recruitment agency's ability to conduct recruitments in a timely manner. A single case can derail a recruitment for years. Although in Tamil Nadu it is relatively rare for this to happen in the PSC's Group Recruitment, in other states this is a significant enough problem that the Supreme Court of India felt the need to make the following comments in its judgment in *Ran Vijay Singh v. The State of Uttar Pradesh*, (2018) 2 SCC 357:

> *It is rather unfortunate that despite several decisions of this Court, some of which have been discussed above, there is interference by the courts in the result of examinations. This places the examination authorities in an unenviable position where they are under scrutiny and not the candidates. Additionally, a massive and sometimes prolonged examination exercise concludes with an air of uncertainty*
>
> *[...]*
>
> *The present appeals are a classic example of the consequence of such interference where there is* ***no finality to the result of the examinations even after a lapse of eight years*** *[emphasis added]. Apart from the examination authorities even the candidates are left wondering about the certainty or otherwise of the result of the examination – whether they have passed or not; whether their result will be approved or disapproved by the court; whether they will get admission in a college or university or not; and whether they will get recruited or not. This unsatisfactory situation does not work to anybody's advantage and such*

> *a state of uncertainty results in confusion being worse confounded. The overall and larger impact of all this is that public interest suffers.*

There is little recruitment agencies can do to prevent a recruitment from being delayed once it gets stuck in court. But one way that recruitment agencies can proactively avoid these situations is to create exams with less score compression.

By setting exams with a wider distribution of test scores, recruitment agencies may be able to reduce the risk of a petition delaying the final selection outcome. In its rulings, courts often assess whether addressing the petitioner's grievance would cause them to fall within the "zone of consideration" for a post.[62] If the score gap is a few points, then challenging any given rule or procedure could potentially put a petitioner in the zone of consideration. When the score gap is wider, petitioners would likely need to raise more substantive challenges to the selection process to receive serious consideration. This limits the likelihood that selection outcomes will be disputed.

[62] See, for example, the Madras High Court's ruling in *J. Menaka vs. The Tamil Nadu Public Service Commission*, Dated 08.09.2022

A more precise exam can potentially help build trust in the recruitment process

At present, many candidates doubt the integrity of the selection process. In a pilot survey of 221 candidates, I asked respondents to provide recommendations for how to fix any shortcomings that TNPSC might have. A common complaint was that TNPSC should be less corrupt.[63] Some of these types of responses include:

[63] In the same survey, 102 out of 221 respondents (46%) said that it was either "maybe" or "definitely" the case that coaching centers were bribing TNPSC officials to get better placement for their students.

- Response 1: "வயதை வீணாக்கி, உழைத்து, படித்து எவனோ ஒருவர் லஞ்சம் குடுத்து வேலையை பெறுவது வெட்க கேடானது"

  [Translation: It is a shame that someone wastes their age, works, studies and the other one gets a job just giving bribe]

- Response 2: "தேர்வைணவர்களை பணியில் சேர்க்க வேண்டும் நேர்மையான முறையில் அதிகரா தூஸ்ப்பிரயோகம் இருக்க கூடாது அரசியல் வாதிகள் தலையீடு இருக்க கூடாது இன்னும் பல..."

  [Translation: Candidates should be recruited in an honest manner, there should not be abuse of power, there should not be interference of politicians, etc.]

- Response 3: "Maximum fraud activities going on exam board please don't do that activities" *[sic]*

This mistrust is not without reason. In the past 10 to 12 years, there have been several verified corruption scandals in TNPSC, and in other Public Service Commissions around the country. In light of these scandals, even if Public Service Commissions conduct recruitments honestly, they face a highly skeptical public.

Still, because trust in recruitment is crucial to establishing the legitimacy of the bureaucracy, it is vital that Public Service Commissions find ways of rebuilding that trust.

Decompressing the score distribution may be one way to do so. When there are only a few points separating candidates who are selected from those who are not, it can be hard for not-selected candidates to accept the results of the exam. After all, from these candidates' point of view, the exam results prove that they are of nearly equal ability to those who were selected. To cope with the disappointment, candidates may be more inclined to believe they were left out because of foul play.[64]

[64]Kuipers (2022) shows that, in the Indonesian civil services exam, candidates who *just* missed the cutoff are more likely to say there is corruption in the selection process than those who are just selected. He then shows this effect is not driven just by changes in attitudes of selected candidates, but also changes in the attitudes of not-selected candidates.

This is where investing in exam quality can pay off. The tools that are used to measure question quality also allow recruitment agencies to standardize question difficulty (and thus the distribution of final test scores). This makes it possible for recruitment agencies to develop multiple versions of the same test and be reasonable confident that the final distribution of scores will be similar for each version. That same technology can be applied to generate practice tests that can closely mimic the real one.

Imagine, instead, that candidates had access to practice tests that provided reliable signals of how well they would do on the real exam, and they could see for themselves that their exam score lined up with how well they had been doing. Moreover, suppose that these tests were sufficiently precise that candidates tend to get the same outcome across repeat iterations, e.g. if they take ten different practice tests they find that they obtain roughly the same score each time. Under these conditions, it would be much harder to disbelieve the selection results—and it would be much easier for recruitment agencies to defend their reputations.

The challenge in realizing this vision is that recruitment agencies need to have a precise method of standardizing exam difficulty. Item Response Theory offers a tried and tested way of doing so.

Exams with more score compression may have higher reapplication rates

It is costly—both for candidates and for society as a whole—for candidates to keep re-applying (see Chapter 2). One of the reasons why candidates keep re-applying is they need to have the satisfaction that there is nothing more that could have been done. As one candidate put it in an interview: *"At least if we gave some four attempts then we can console ourselves that we tried and didn't succeed."*

But why should it take four attempts to come to this conclusion? Can we design the exam in a way that each attempt provides candidates with more information? De-compressing the score distribution would make the result of any given attempt more informative. This, in turn, may reduce the number of attempts candidates would require to come to the conclusion they are unlikely to be selected.

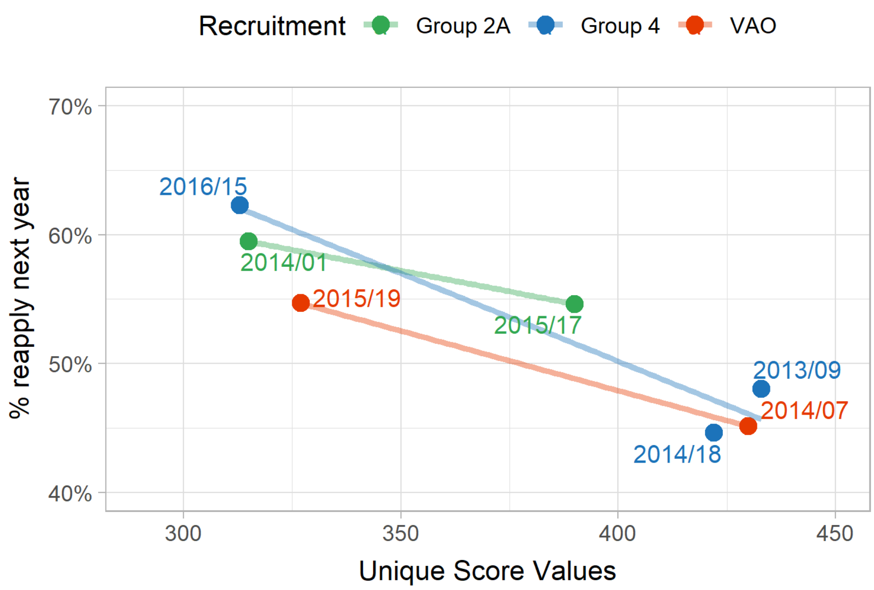
Figure 5.2 shows that there is in fact a stark correlation between score compression and the re-application rate. The horizontal axis measures the number of unique scores in the final distribution; the larger the value, the less compressed the score distribution is.[65] The vertical axis is the fraction of candidates that re-apply in any TNPSC exam in the following year.

[65] To illustrate how this variable is calculated—if there are 10 candidates, 3 of whom scored 240 marks and 6 of whom scored 200 marks, the value on the axis will be 2, because that is the total number of unique score values in the sample.

The relationship is not just driven by variation in the number of applicants. One might worry that we are just capturing the fact that when more candidates apply, the score distribution tends to get more compressed, and the marginal applicant may be less committed to re-applying. However, in some cases, exams that have more applicants also have *less* score compression. For example, in Group 2A, the 2015 exam had about 36% more applicants than the 2014 exam, but had less score compression.

The figure only plots a correlation, so it is only suggestive of a causal relationship. But it raises the interesting possibility that exam design can reduce excess re-application. This idea merits further consideration.

Figure 5.2: Score Compression and Reapplication Rates



Notes: The horizontal axis is the number of unique values of the score variable in the whole population of applicants. The vertical axis is the fraction of candidates that re-apply in any TNPSC exam in the following fiscal year. The solid lines are regression lines. The sample is restricted to the single-stage exams.

Standardizing exam difficulty is the long-term solution to question paper leakage

It is challenging for recruitment agencies to handle the large application volume that they currently do. TNPSC, for example, can receive over 15 lakh applications for a single recruitment (see Chapter 1). In order to conduct a recruitment for so many candidates at the same time, the government needs to set up thousands of exam centers across the state.

In doing so, the government spreads its resources thin. Each TNPSC staff member becomes responsible for overseeing a proportionately larger number of exam centers. Moreover, a large number of other government officials need to get involved in securing the exam center, and transporting the question papers and answer sheets to and from headquarters. The more people get involved, the more vulnerabilities in exam security are created. As a result, recruitment agencies struggle to prevent question papers from leaking.

If recruitment agencies no longer had to conduct the recruitment for all candidates at the same time, it would open up a number of new options for managing the exam process. In particular, this would allow recruitment agencies to conduct exams through Computer Based Tests. The key advantage of the Computer Based Test is that the questions need not leave headquarters until the moment the exam is live. The catch is, there is no way the government can procure enough computers to conduct Computer Based Tests for over 15 lakh applicants simultaneously. As a result, recruitments that are conducted simultaneously cannot adopt this technology.

The standard solution to the problem is to split the applicant pool into batches, and test them on separate versions of the test. This approach has been tried in several prominent recruitments, including the Railway Recruitment Board's Group D exam. To adjust for the fact that the difficulty of the exams across sessions is not comparable, Indian recruitment authorities generally rely on linear equivalence scales. These equivalence scales are only valid under strong assumptions about the distribution of test scores. In the absence of a well-defined process of pre-testing and validating questions before they are fielded, it is extremely risky for a recruitment authority to depend so heavily on these assumptions.

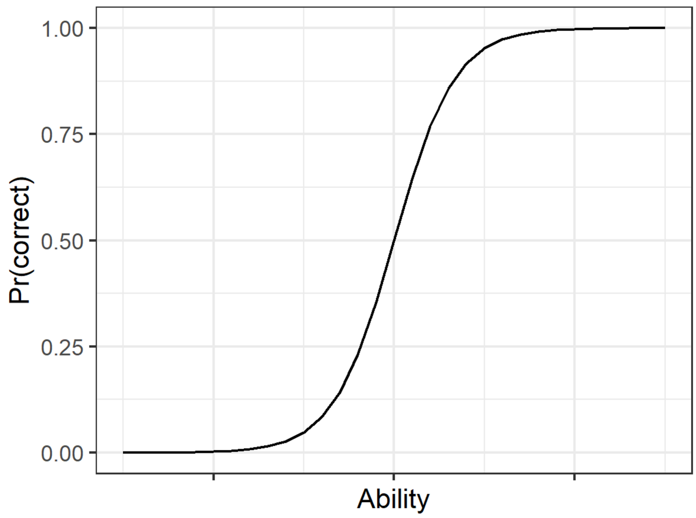
However, as mentioned above, investing in exam quality allows recruitment agencies to standardize the difficulty of the exam. Once this ability is in place, recruitment agencies may no longer have to link exams and notifications at all. For example, candidates could take the exam in small batches on a continuous basis, at their convenience, and obtain a score that is valid for an extended duration of time; in turn, government departments would no longer need to wait up to several years for the recruitment process to complete, but could recruit the next candidate on the merit list whenever a vacancy opens up. This—and many other possibilities—open up once the exams are standardized.

5.3 What makes a good question?

How do we identify good questions?

Item Response Theory is based on the idea that there is a curve, called the ***Item Characteristic Curve (ICC)***, which describes the relationship between a candidate's ability, and the fraction of candidates who answer the question correctly. The curve plotting this relationship usually looks like this:

Figure 5.3: The Item Characteristic Curve



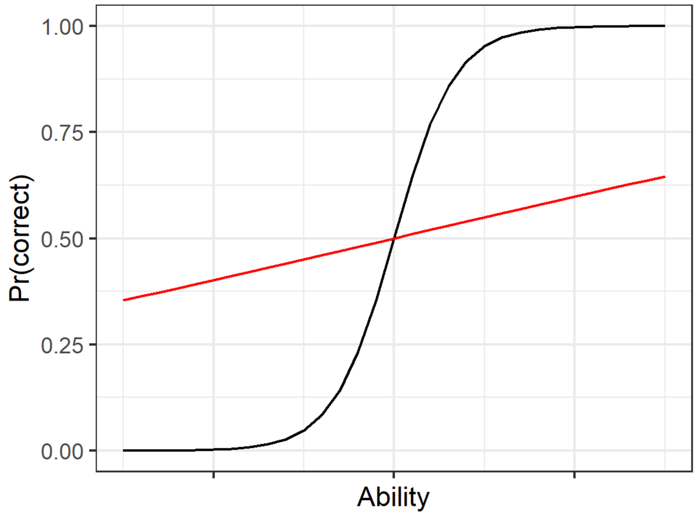
Question quality can be defined in terms of the shape of this curve. I will illustrate how question quality affects the shape of the ICC using some hypothetical examples. In each of the following scenarios, the high quality question is shown in black, and the low quality question is marked in red.

1. Good questions have high discrimination

The discrimination of an exam question is a measure of how sharply the probability of answering correctly increases across the ability distribution. Questions with higher discrimination result in less score compression.

If a question has high discrimination, then the middle part of its ICC will be steep (the black line in Figure 5.4). If it has low discrimination, then this part of the curve will be relatively flat (see the red line). Intuitively, when a question has low discrimination, then observing a correct answer tells us little about their ability, because candidates have similar probabilities of getting a correct answer. In this example, the red line shows that most candidates have a probability of answering correctly that is close to 50%.

Figure 5.4: High vs. Low Discrimination



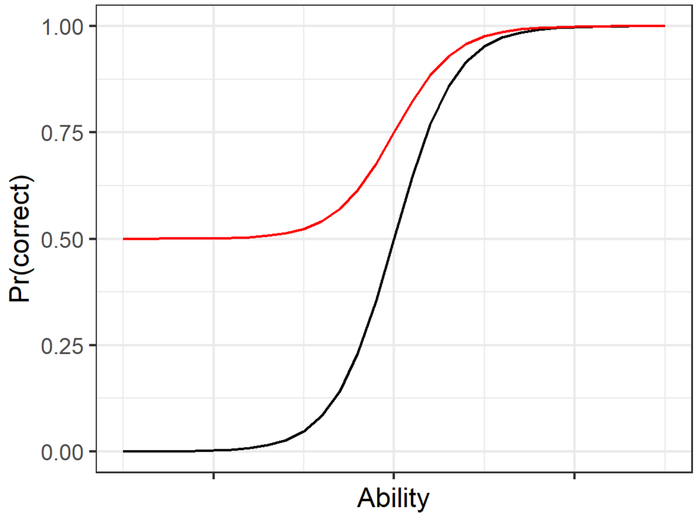
2. The probability of guessing correctly should be low

If the probability of guessing correctly is high, then there is less variation in the probability of answering correctly across the ability distribution. A high guessing probability therefore compresses the score distribution.

We can tell if a question has a high probability of guessing correctly if the left part of the ICC tends towards a number much larger than zero (see, e.g. the red curve in Figure 5.5). In fact, the number that the curve tends to is the probability of guessing correctly (i.e. because that is the probability of answering correctly for candidates that almost surely do not know the answer).

Thus, the black curve plots the ICC for a question where the probability of guessing is close to zero, and the red curve is the ICC plot where the probability of guessing is 50%.

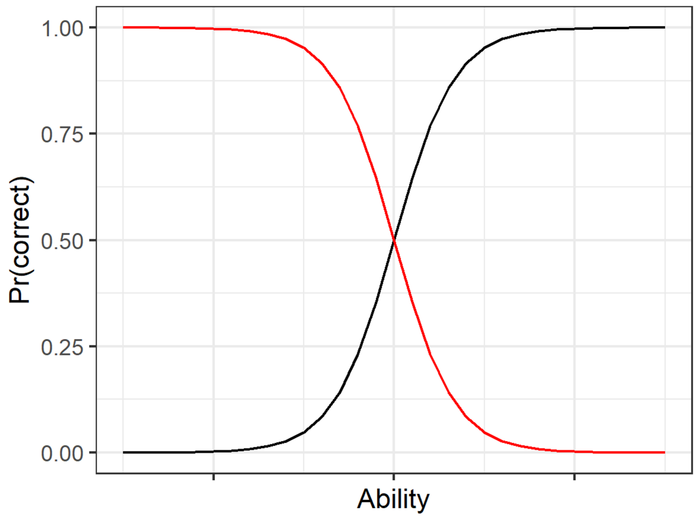
Figure 5.5: High vs. Low Guessing Parameter



3. Questions on the same exam should mostly measure the same underlying ability

ICCs may not slope in the same direction, as in Figure 5.6.

Figure 5.6: An example of inconsistency across questions



If that happens, it means that the exam is measuring distinct traits that tend to be opposed to each other. For example, most questions may measure analytical ability, but one question may measure the capacity to memorize; if the candidates who are good at thinking analytically are also bad at memorizing, then we may end up with the pattern we see in Figure 5.6.

Testing for multiple, opposing traits is not inherently undesirable. For example, the nature of the job may demand that candidates have these opposing traits, even though *on average* they tend not to coincide in the population.

However, the cost of testing for opposing traits is that it compresses the score distribution.

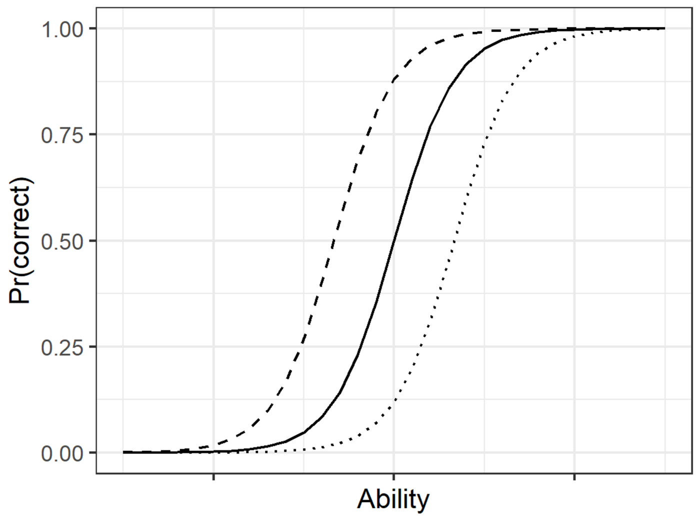
4. The exam questions need to be concentrated at the right level of difficulty

Each individual item distinguishes candidates only at specific points in the ability distribution. If someone answers an easy question incorrectly, chances are their ability is low; but it does not provide us with much information on how high their ability is. Similarly, if someone answers a hard question correctly, we know their ability is likely to be very high; but if they answer incorrectly, we have little information on how low the ability is.[66]

[66] As a concrete example, if a 3rd grade student answers a calculus question correctly, then we know their math knowledge is very high; if they answer incorrectly, we really have no idea how well they understand more grade-appropriate concepts.

Figure 5.7 shows how ICCs vary with question difficulty.

Figure 5.7: Item Characteristic Curves for Questions of Varying Difficulty



In order to ensure that test scores are sufficiently spread out, we need to make sure there are enough questions of the right difficulty. The more questions there are at a given level of difficulty, the more spread out that test scores will be among candidates who are just able to answer those questions. Easy questions have little value, because once candidates know they are far away from the cutoff, there is no need to separate them further.

In what ways can questions improve?

If we have data from a multiple choice test on which questions candidates answered correctly and which ones they answered incorrectly we can estimate ICCs for each question on the test. To do so, we fit a statistical model to the data. (Appendix B.4 provides more details on the estimation procedure).

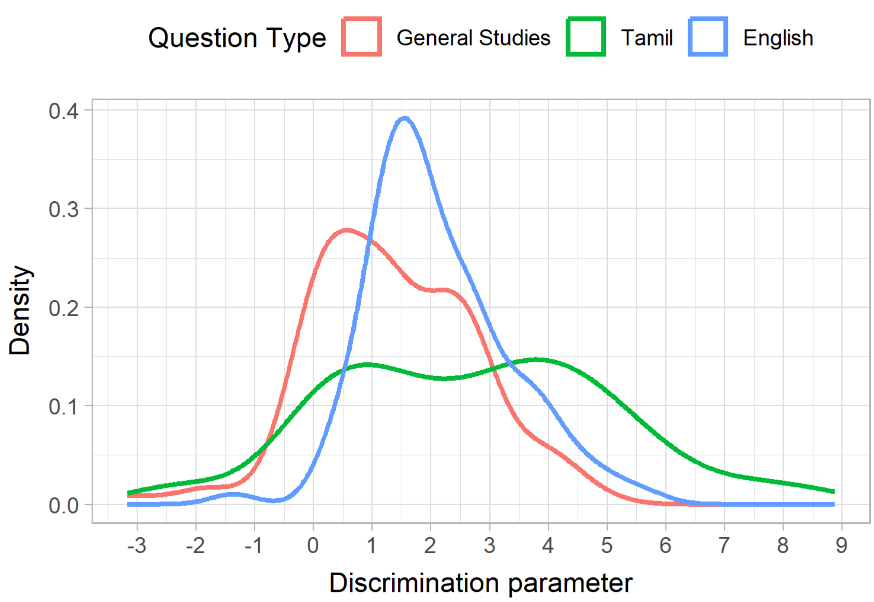
In this section, I illustrate how an analysis of exam quality based in Item Response Theory can identify potential areas of improvement. For this illustration, I use one of the Group 4 exams as an example.

Here is what I find:

The discriminatory power of questions could be improved.

Experts consider questions to be high quality when they have discrimination coefficients of 0.8 or above (Baker, 2001; De Ayala, 2013). About 23% of the questions on this exam currently do not meet these standards, and would therefore benefit from revision.

Figure 5.8: Variation in Item Discrimination



There are substantial differences in the discrimination parameter across the different components of the exam (see Figure 5.8 below). Questions from the English section of the exam are generally high quality, with only 8% not meeting the benchmark. Meanwhile, for Tamil and General Studies, only 25% and 37% of questions do not meet this benchmark, respectively.

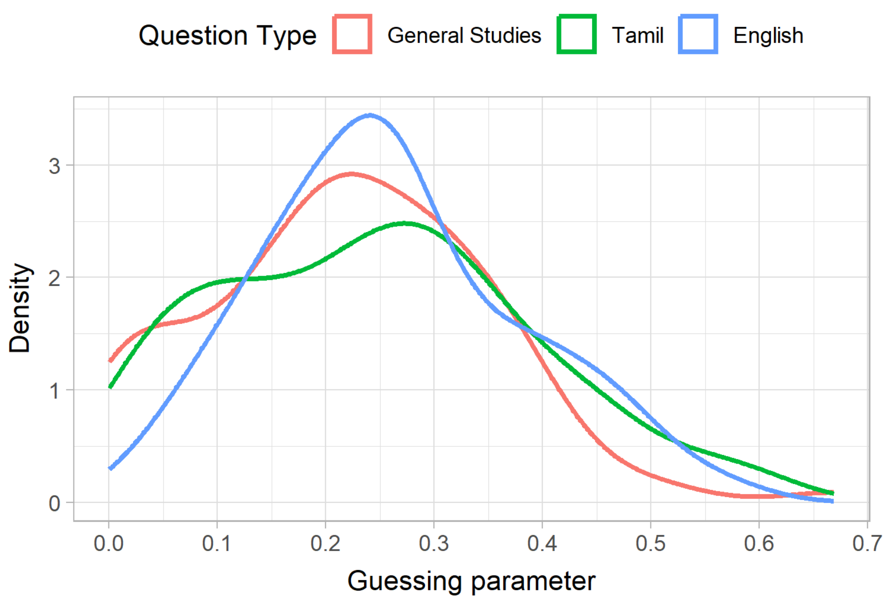
Guessing the correct answer can be made harder

In the Group 4 exam, each question is presented with four possible responses. Thus, the baseline probability of guessing correctly is 25%. If the guessing rate for a particular question is higher, that means that some of the wrong options tend to be easily discarded by applicants. If the guessing rate is lower, it means that one of the wrong answers looks particularly appealing to candidates who do not know the correct answer.

Figure 5.9 plots the distribution of the probability of guessing the correct answer for each question.

In this exam, the median question has a guessing parameter close to 0.25, i.e. candidates have a 25% chance of guessing the correct answer for the typical question. This suggests that most questions are doing a fairly good job of presenting plausible incorrect responses. However, in 46% of questions the guessing parameter exceeds 0.25. These questions can be revised so that the incorrect answers are more compelling.

Figure 5.9: Variation in Item Guessing Rates



Exam questions can be made more consistent with each other.

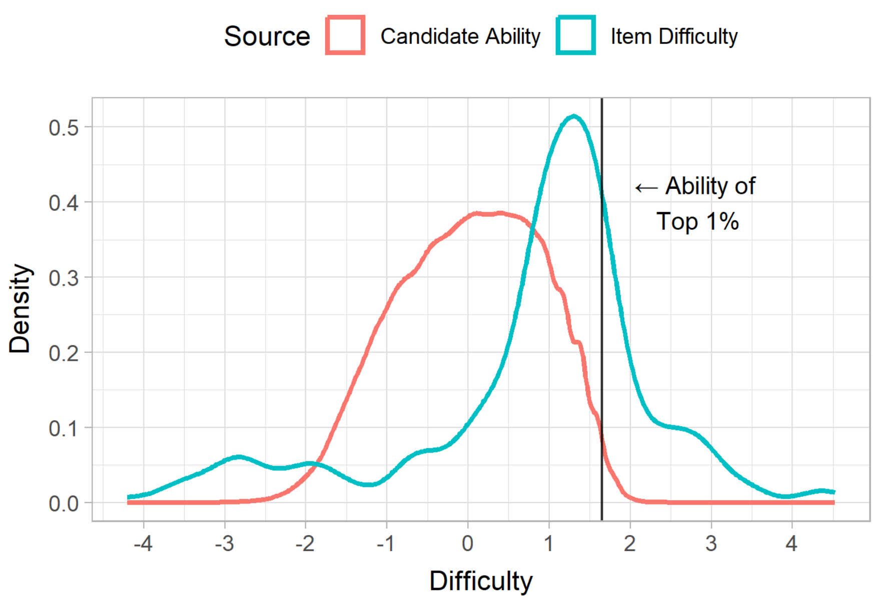
Note in Figure 5.8 that about 8% of questions have a negative discrimination parameter. These questions should be more closely examined to determine whether they are intended to test separate concepts from the rest of the section, or whether they can be improved.

Question difficulty is not a problem. Formulating difficult questions that top-scoring candidates can consistently answer correctly is.

Figure 5.10 plots the variation in item difficulty (as defined in the IRT model), and compares it to the distribution of candidate ability.

It turns out that the difficulty level is already well-targeted. The more questions there are at a particular ability level, the better able the test is to separate candidates at that level from everyone else. The fact that most questions are concentrated in the top part of the distribution, close to the ability of the top 1% of candidates (marked by the vertical line), should, all else equal, *help* reduce compression at the top of the distribution.

Figure 5.10: Variation in Item Difficulty Relative to Candidate Ability



Notes: For clarity, about 10% of items with extreme values have been removed.

Why, then, do these questions generate very little separation in scores at the top of the distribution? The answer is that their discrimination is low.[67] Low discrimination tells us that top-scoring candidates are not able to *consistently* answer difficult questions well. Instead, sometimes they answer correctly, sometimes not; and a fair number of times—for questions where the discrimination parameter is negative—the top-scoring candidates are *less* likely to get the right answer. Put another way, not enough lower ability candidates get the difficult questions *consistently* wrong that these questions are able to generate much separation.

[67] One symptom of this is that the fraction of candidates answering the item correctly is more strongly correlated with the item's discrimination parameter rather than its difficulty parameter.

What kind of changes to the questions would produce the desired effect on the underlying parameters?

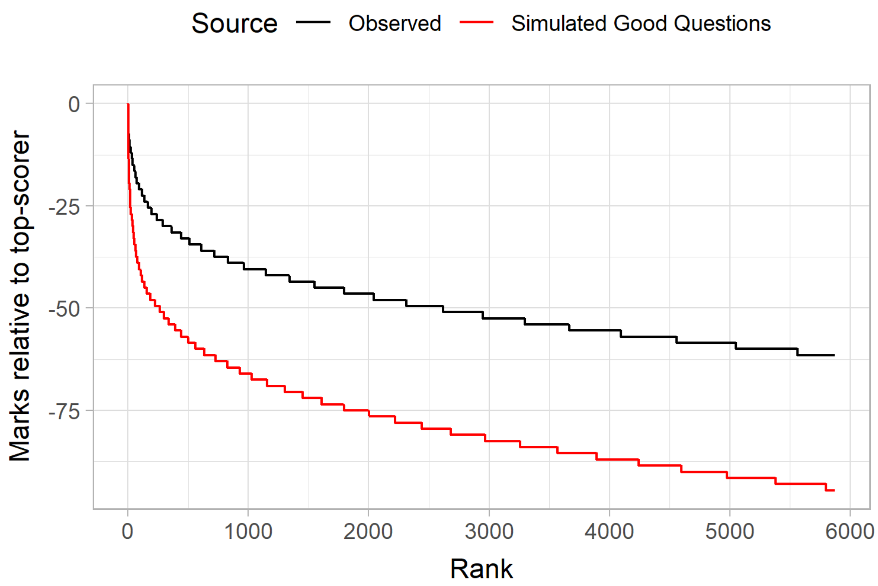
A nuanced diagnosis of question quality allows us to make educated guesses about the right direction for how to try to improve them. For example, the fact that it is discrimination that matters and not difficulty suggests that exam designers can try focusing on building questions based on an underlying, consistent set of concepts. That way, candidates who know the concepts will be able to answer all the questions that use a similar concept correctly; and those who do not know will not. To test this hypothesis, a recruitment agency could try developing questions along these lines, piloting them with a sample population, and observing whether there is an improvement on these key metrics of question quality.

By how much can question quality affect score compression?

Suppose that a recruitment agency implements the ideas described above. By how much is it theoretically possible for recruitment agencies to decrease score compression?

Figure 5.11 presents the results of a simulation exercise that addresses this question. The figure shows the distribution of test scores at the top of the distribution, both for the marks that are observed, and for a hypothetical set of marks that would be generated if the questions were highly optimized.

Figure 5.11: By how much can good questions reduce score compression?



Notes: See Appendix B.4 for details on how the simulated curve was generated.

The potential gains to improving question quality are large. The fact that the red curve is everywhere below and steeper than the black curve means that scores are more spread out. In this simulation, the number of unique values among the top-scoring candidates increases by 48%, and the gap between the highest and lowest score in this range increases from 61.5 to 97.5.

5.4 Discussion

What would it take to improve exam quality?

What kind of resources do recruitment agencies need to work on improving exam quality?

In the short-run, I suspect that recruitment agencies will be able to make dramatic improvements in question quality as long as they have access to someone who can analyze data—even without the help of professional psychometricians.

Although all-purpose data analysts will not have the same level of expertise, there are still many low-hanging fruits that data analysts can help recruitment agencies obtain. Their main value will not be in subtle optimization, but rather in identifying some of the worst performing questions and making sure that they are not fielded in a live recruitment.

In the long-run, it is possible that India will increase its domestic supply of psychometricians. Some of this capacity may come from the National Testing Agency, a Union Government body that has already partnered with leading experts and is specifically mandated with providing training and technical assistance.

How can recruitment agencies use data to improve question quality?

There are two main ways that recruitment agencies use the outputs of an analysis of question quality.

First, recruitment agencies typically rely on a pool of question-setters who are subject matter experts. With a detailed analysis of question quality, recruitment agencies can identify whether there are any question-setters who tend to systematically provide low quality questions, and rely on them less in future examinations.

Second, recruitment agencies can use this information to identify common patterns in questions that worked well in past exams. Recruitment agencies can then use this understanding to develop new questions. For example, recruitment agencies can communicate this understanding to question setters by publishing guidelines and recommended formats that questions should take. Alternatively, there may be ways in which questions can be generated by computers. For example, one could use an algorithmic model to generate compelling wrong answers.[68] With systematic research in this area, it may be possible to fully automate some part of question development.

[68]See, e.g., Murugan and Balasundaram (2021) for a discussion of how this can be done for Tamil vocabulary tests

What would it take to standardize exam difficulty?

This is a challenging goal. Standardizing exam difficulty requires knowing how difficult each question will be *before* it is fielded in a live exam. But often the only way of knowing how difficult a question will be with a meaningful amount of precision is to collect data on how well candidates are able to answer it. The problem is that releasing questions in advance threatens to compromise the integrity or usefulness of the questions in future exams.

There are several ways that testing agencies usually get around this problem. Recruitment agencies would likely need guidance from experts with past experience in order to successfully adopt any of these strategies for question development:

1. Flood the market with questions If candidates know that a specific set of 200 questions will appear on the next exam, then the integrity of that exam is compromised. But if candidates only know that some set of 200 questions from a bank of 20,000 will appear on the next exam, then the integrity of the exam is no longer at risk. After all, if they manage to study for all of these, then they have demonstrated that they possess the knowledge that the exam assesses. The key is to maintain enough uncertainty around which questions will be chosen so that it would be difficult for candidates to use the availability of prior questions to their advantage.

To do this, recruitment agencies would need to invest heavily in continuous question development, rather than developing questions only for specific, upcoming exams.

This operation also requires recruitment agencies to figure out how to select who will test the questions. One way of doing this is to bundle it with an outreach effort aimed at making practice tests more easily available to candidates. For example, recruitment agencies can partner with schools and universities to offer practice tests. The practice tests can recycle old test questions, while also testing experimental questions.[69] This data can then be sent back to recruitment agencies for analysis.

[69] There is a lot of value in mixing old questions with new ones because that allows recruitment agencies to adjust for differences in the characteristics of the test population with the population that actually appear for live exams.

Recruitment agencies would need to have a strategy for avoiding giving candidates who participate in the practice tests an unfair advantage. One possibility is to use small test populations and wait until the database of piloted questions is sufficiently large. How exactly this should be done would require detailed consultations with experts who have managed such assessments before.

2. Test new questions in live recruitments Recruitment agencies can introduce a set of questions that are non-scoring. These questions can be randomly distributed throughout the exam, and candidates would not be told which questions are non-scoring (so that they take all questions seriously).

The main advantage of this approach is that it allows recruitment agencies to innovate without requiring them to make as many changes to their existing operations. The main catch is that this format forces the speed of innovation to be slower, since the new questions cannot be too many, or too different from the rest of the exam. It therefore requires some strategic foresight to figure out what kinds of questions are best tested in this manner.

Conclusion

To conclude, I present some of the common themes that have emerged from the previous chapters.

The value of data

Recruitment agencies should consider employing data analysts

As we have seen, there is a wealth of information contained in the data that recruitment agencies have. For example, this data can be used to:

- Inform the government of changing conditions in the labor market (see Chapter 1)
- Communicate vital information to candidates about their progress through the selection process (see Chapters 3 and 4)
- Improve exam quality (see Chapter 5)

These initiatives have the potential to not only help recruitment agencies improve their own processes, but they can also have impacts in the broader labor market as well.

However, in order for recruitment agencies to make full use of this information, they need to employ data analysts who have the necessary skills. Excel is not sufficient. This kind of data will need to be analyzed using a programming language such as R, Stata, or Python.

People with these kinds of skills are generally rare in government departments. However, data analytics units are starting to become increasingly common, even at the state level. For now this trend has largely escaped recruitment agencies, but given the potential benefits it may be worthwhile for them to start building this capacity internally.

Household surveys should include questions that identify candidates preparing for competitive exams

An individual recruitment agency only observes the candidates who are preparing for their own exams. But candidates routinely apply for exams *across* recruitment agencies. In order to get a reliable picture of total application behavior, we need data that provides information on exam participation for representative samples. The best way of doing that in India is through a government-run household survey, such as the Periodic Labour Force Survey (PLFS).

Making public sector recruitment more effective

Recruitment agencies should try to educate candidates about the selection process

The Candidate Survey revealed that many TNPSC candidates remain severely misinformed about basic facts about the recruitment process. For example: most have very little knowledge of what the cutoff will be (see Figure 3.14); they do not realize how informative their past test scores are of their future selection probabilities (see Figure 3.16); and they dramatically under-estimate the number of attempts that selected candidates usually make (see Figure 4.2).

Recruitment agencies have data that, if shared with the public, could help correct these misunderstandings. By better aligning expectations with reality, recruitment agencies may be able to win candidates' trust, reduce excess applications, and potentially even improve candidates' labor market outcomes. The policy challenge lies in figuring out how to communicate the difficult truths about the recruitment process in a sensitive manner.

Lessons for Labor Market Policy

To address the problem of educated unemployment, focus on supply, not just demand

The Indian government's approach to the problem of educated unemployment largely focuses on the demand side of the market. There is a heavy emphasis on the idea that young people are not employed because they are not "employable," and that their skill levels must increase before it is worthwhile for firms to put them on payroll.

At the same time, we have seen evidence throughout the report that complicates this narrative. Educated young people who are currently unemployed often put little effort into job search, and many are reluctant to join the private sector. A surprisingly large share indicate that they would not be working even if they were not preparing for competitive exams.

This is especially true for women, who are concerned about finding work that allows them to also manage their household responsibilities.

Moreover, it is not the case that they are unable to afford investment in skills, or unwilling to do so. Rather, the skills that many educated young people would rather focus on are the ones that are valued by public sector recruitment agencies.

Together, this evidence suggests that reducing unemployment is as much a problem of supply as it is a problem of demand. Not only do there need to be enough firms demanding work, but the jobs they offer need to be sufficiently desirable.

Of course, there may not be a clean separation between supply and demand. For example, when there are more firms competing for workers, they can compete for workers by upgrading the quality of the jobs they offer. But thinking about unemployment in terms of supply suggests some less explored avenues for policy intervention that may be worth considering, including: curbing false job advertisements in the private sector; focusing on job quality instead of quantity; and reducing some of the sources of monopsony power in the market that keep private sector wages low.[70]

[70] Roughly speaking, monopsony power refers to the fact that firms could afford to pay workers more without threatening their ability to remain profitable, but can choose not to because workers' labor supply is "captive" in some way. Brooks et al. (2021) provides evidence that monopsony power in India can be quite high.

Recruitment agencies should be mindful of how their policy decisions affect the labor market.

As long as government jobs are as attractive as they currently are—and this seems like it will continue to be the case for the near future—recruitment policy will likely affect wages, employment, and skill development in the rest of the economy. In a way, there are interesting parallels between the role of central banks in credit markets and the role of recruitment agencies in the labor market: both are agencies that are often independent of the government and can shape how the market evolves through their policy decisions.

This is both an opportunity and a challenge. The opportunity is that recruitment policy can potentially drive important reforms, especially in stubborn and challenging problems such as educated youth unemployment. The challenge is that we currently have very little evidence on how to use recruitment policy to improve labor market outcomes.

Future Directions

We have much more to learn

In many parts of this report, the best I could offer was either suggestive evidence or open hypothesizing. These hunches may or may not be true. With time, I hope that our understanding will come to sit on firmer foundations, and that we can marshal that understanding to develop more effective recruitment and labor market policies.

References

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

Arellano-Bover, Jaime (forthcoming), "Career consequences of firm heterogeneity for young workers: First job and firm size." *Journal of Labor Economics*.

Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor (2016), "Does affirmative action work? caste, gender, college quality, and academic success in india." *American Economic Review*, 106, 1495–1521.

Baker, Frank B (2001), *The basics of item response theory*. ERIC.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan (2010), "Affirmative action in education: Evidence from engineering college admissions in india." *Journal of Public Economics*, 94, 16–29.

Brooks, Wyatt J, Joseph P Kaboski, Yao Amber Li, and Wei Qian (2021), "Exploitation of labor? classical monopsony power and labor's share." *Journal of Development Economics*, 150, 102627.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin (2022), "Job search and hiring with limited information about workseekers' skills." *American Economic Review*, 112, 3547–83.

Chalmers, R Philip (2012), "mirt: A multidimensional item response theory package for the r environment." *Journal of statistical Software*, 48, 1–29.

Colonnelli, Emanuele, Joacim Tåg, Michael Webb, and Stefanie Wolter (2018), "A cross-country comparison of dynamics in the large firm wage premium." In *AEA Papers and Proceedings*, volume 108, 323–27.

Dahis, Ricardo, Laura Schiavon, and Thiago Scot (2020), "Selecting top bureaucrats: Admission exams and performance in brazil." *Available at SSRN 3584725*.

De Ayala, Rafael Jaime (2013), *The theory and practice of item response theory*. Guilford Publications.

Ferrara, Eliana La, Alberto Chong, and Suzanne Duryea (2012), "Soap operas and fertility: Evidence from brazil." *American Economic Journal: Applied Economics*, 4, 1–31.

Finan, Frederico, Benjamin A Olken, and Rohini Pande (2017), "The personnel economics of the developing state." *Handbook of economic field experiments*, 2, 467–514.

Friedberg, Rina, Julie Tibshirani, Susan Athey, and Stefan Wager (2020), "Local linear forests." *Journal of Computational and Graphical Statistics*, 30, 503–517.

Goldin, Claudia and Lawrence F Katz (2011), "The cost of workplace flexibility for high-powered professionals." *The Annals of the American Academy of Political and Social Science*, 638, 45–67.

Hjort, Jonas, Xuan Li, and Heather Sarsons (2020), "Across-country wage compression in multinationals." Technical report, National Bureau of Economic Research.

Imbert, Clement and John Papp (2015), "Labor market effects of social programs: Evidence from india's employment guarantee." *American Economic Journal: Applied Economics*, 7, 233–63.

Jeffrey, Craig (2010), *Timepass: Youth, class, and the politics of waiting in India*. Stanford University Press.

Jha, Mrinalini and Amit Basole (2023), "Labour incomes in india: A comparison of two national household surveys." *The Indian Journal of Labour Economics*, 1–21.

Khanna, Gaurav and Nicolas Morales (2021), "The it boom and other unintended consequences of chasing the american dream." Unpublished manuscript.

Kuipers, Nicholas (2022), "Failing the test: The countervailing attitudinal effects of civil service examinations." *American Political Science Review*, 1–18.

La Ferrara, Eliana (2016), "Mass media and social change: Can we use television to fight poverty?" *Journal of the European Economic Association*, 14, 791–827.

Mangal, Kunal (2022a), "How much are government jobs in developing countries worth?" Unpublished manuscript.

Mangal, Kunal (2022b), "The long-run costs of highly competitive exams for government jobs." Unpublished manuscript.

Manning, Alan and Barbara Petrongolo (2008), "The part-time pay penalty for women in britain." *The economic journal*, 118, F28–F51.

Ministry of Skill Development and Entrepreneurship, Government of Inida (2021), "Annual Report 2021-2022." Technical report.

Mueller, Andreas I, Johannes Spinnewijn, and Giorgio Topa (2021), "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias." *American Economic Review*, 111, 324–363.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar (2016), "Building state capacity: Evidence from biometric smartcards in india." *American Economic Review*, 106, 2895–2929.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar (2017), "General equilibrium effects of (improving) public employment programs: Experimental evidence from india." Technical report, National Bureau of Economic Research.

Murugan, Shanthi and SR Balasundaram (2021), "Affix-based distractor generation for tamil multiple choice questions using neural word embedding." *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 13.

Narayanan, Sudha and Upasak Das (2014), "Women participation and rationing in the employment guarantee scheme." *Economic and Political Weekly*, 46–53.

Pandey, Nidhi Gyan (2021), "Women's Workforce Participation in India: Statewise Trends." Technical report, IWWAGE – Lead at KREA University.

Poonam, Snigdha (2021), "India has a fake-jobs problem." *The Atlantic*.

Riley, Emma (2019), "Role models in movies: the impact of queen of katwe on students' educational attainment." *The Review of Economics and Statistics*, 1–48.

Sekhri, Sheetal (2020), "Prestige matters: Wage premium and value addition in elite colleges." *American Economic Journal: Applied Economics*, 12, 207–25.

Special Correspondent (2022), "Over 18.5 lakh candidates appear for tnpsc group iv exam." *The Hindu*, URL `https://www.thehindu.com/news/cities/chennai/over-185-lakh-candidates-appear-for-tnpsc-group-iv-exam/article65678612.ece`.

Staff Selection Commission (2021), "SSC Annual Report 2020-2021." Technical report.

Tamil Nadu Skill Development Corporation (2019), "Skilling for the Future: Skill Gap Assessment and Action Plan for Tamil Nadu." Technical report.

TN Teacher Recruitment Board (2022), "Tamil nadu teacher eligibility test." `https://www.trb.tnschools.gov.in/demo/tntet?language=LG-1&status=Active`. Accessed: 2012-01-18.

Vardhan, Anand (2017), "Coaching centres: The ias vending machines of delhi." *Newslaundry*, URL `https://www.newslaundry.com/2017/03/28/coaching-centres-the-ias-vending-machines-of-delhi`.

Vijay Kumar, S. (2020), "Tamil nadu police recruitment hit by certificate scam; 1000 candidates involved." *The Hindu*, URL `https://www.thehindu.com/news/national/tamil-nadu/tamil-nadu-police-recruitment-hit-by-certificate-scam-1000-candidates-involved/article30773019.ece`.

Wager, Stefan and Susan Athey (2018), "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association*, 113, 1228–1242.

Xu, Guo, Marianne Bertrand, and Robin Burgess (2021), "Organization of the state: Home assignment and bureaucrat performance." *The Journal of Law, Economics, and Organization*.

Xu, Huayu and Achyuta Adhvaryu (2022), "The human capital effects of access to elite jobs." Unpublished manuscript.

Appendices

A Data

A.1 Administrative Data

Included Recruitments This dataset includes the following Group Recruitments:

FY	Recruitments	Level
2013	2013/09	Group 4
	2013/14	Group 2
	2013/17	Group 1
2014	2014/01	Group 2A
	2014/07	VAO
	2014/18	Group 4
2015	2015/07	Group 2
	2015/09	Group 1
	2015/17	Group 2A
	2015/19	VAO

FY	Recruitments	Level
2016	2016/15	Group 4
	2016/19	Group 1
2017	2017/10	Group 2A
	2017/23	Group 4
2018	2018/15	Group 2
2019	2019/01	Group 1
	2019/19	Group 4

Matching Candidates Across Recruitments There is no consistent unique identifier for candidates across exams in the sample window. To solve this problem, I generated an identifier known as a Hash ID. The Hash ID is an encrypted version of the candidate's name and his or her parent's name. The Hash ID is reliably unique within recruitments and across recruitments.[71] I find that the Hash ID uniquely identifies candidates within recruitments in over 99% of cases.

[71] Candidates have an incentive to spell their name consistently in their application form because TNPSC will flag their application if it does not match the name written in their credentials.

A.2 The Candidate Survey

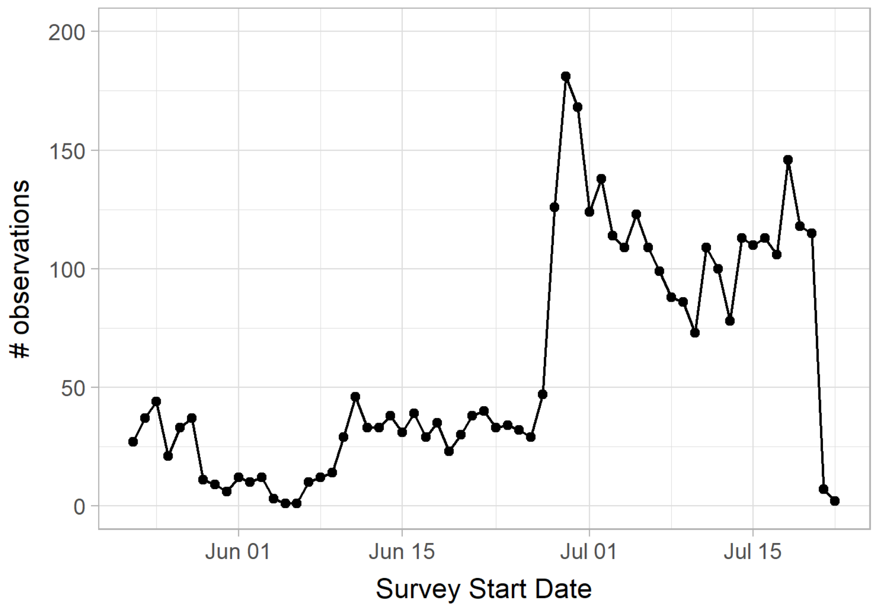
Respondent Sourcing Most candidates were obtained through Facebook advertising, though we did also obtain a small number of respondents through word-of-mouth.

Table A.1: How respondents were sourced for the Candidate Survey

Source	N	% of sample
Facebook	3450	96.5
Word of Mouth	74	2.1
Colleges / Coaching Centers	48	1.3
Telegram	2	0.1

Timing In the baseline survey, responses were collected from between May 23rd, 2022 and July 22nd, 2022. Figure A.1 shows the variation in the number of responses received over time, relative to the time of the exam. About 33% of responses come from within 2 weeks of the exam, and about 78% come from within 1 month of the exam.

Figure A.1: Response Timing in the Candidate Survey



Data Quality Since the survey was online and anonymous, data quality is a potential concern. We implemented several procedures to verify that responses were genuine and help respondents fill out usable data.

First, we called all people who participated in the survey by phone. During the phone conversation, we asked them to verify their name, their date of birth, and their district. If the person was unreachable or they gave incorrect information, then they were marked as an unverified response. In total, 73% of responses were verified. The main reason that we were not able to verify the rest is that they were not comfortable providing their contact information.

Second, we only provided the gift card to candidates who completed the phone verification with us. This policy deterred people from completing the survey multiple times for the sake of obtaining multiple gift cards. In case respondents submitted the same contact information multiple times, the duplicate responses are dropped.

Third, we designed the form in a way that implemented checks and minimized potential errors. For example, when respondents were filling out large numerical values (e.g. in response to, how many candidates do you think will be selected?), the form would display the number they were typing in words.

Follow-up Response Rates We randomized the order in which we called respondents for the Follow-up survey. We were successfully able to complete the survey 96% of our sample.

A.3 Other Data Used in this Report

Tamil Nadu State Board Exam Marks This data is from the Tamil Nadu Directorate of Government Examinations for the 2021-2022 school year.

The Naan Mudhalvan Career Guidance Survey This data is from the Tamil Nadu Department of School Education. In the Fall of 2022, the Department conducted a survey with nearly all 12th standard students about their future career goals. The survey was administered in the school computer labs, and students filled out the responses on their own.

CMIE Consumer Household Pyramids This data is a panel of households, sampled every four months. The survey provides detailed information on household demographics, occupations, employment status, and earnings.

Master Company Registration Data The Ministry of Corporate Affairs provides access to a 2015 copy of its master database of registered companies by state. This is accessible at: `https://www.mca.gov.in/MinistryV2/archiveofmasterdatadetails.html`.

B Estimation

B.1 Estimating the Probability of Selection

This section provides additional details on how the distribution of selection probabilities across candidates was estimated in Section 3.2.

How does it work? I use a random forest algorithm to estimate the probability of selection.

The intuition for this approach is the following: If there is heterogeneity in the probability of selection, then different subgroups of candidates will have different selection rates. For example, we may find that candidates from certain districts tend to have higher selection rates; or that candidates with certain types of degrees have higher selection rates. To uncover this heterogeneity, I supply the algorithm with a set of variables to use for prediction. The algorithm then uses these variables to keeps splitting the sample into smaller and smaller sub-groups. The predicted probability of selection is the fraction of candidates who are selected in the final-subgroups.

The strength of this algorithm thus depends on the quality of the predictors supplied.

Estimation I use the following variables to predict selection:

Exam Level	Predictors Used
Group 1	Communal category, marital status, highest education completed, birth district, pincode of permanent address, whether the candidate is a current government employee, variables related to special reservation categories (widow, ex-service, blind, deaf, orthopeadically disabled, has multiple disabilities), whether 10th standard was completed in Tamil, whether 12th standard was completed in Tamil, whether UG was completed in Tamil, 10th standard board, undergraduate degree, undergraduate major, postgraduate degree, whether they have an integrated postgraduate degree, gender, age, registration date

Notes: Table continued on the next page.

Exam Level	Predictors Used
Group 4	Communal category, marital status, highest education completed, birth district, pincode of permanent address, whether the candidate is a current government employee, variables related to special reservation categories (widow, ex-service, blind, deaf, orthopeadically disabled, has multiple disabilities), undergraduate degree, undergraduate major, postgraduate degree, gender, whether they have an integrated postgraduate degree, age, registration date

Validation How well does the model fit the data?

To evaluate model fit, I train the data on half of the sample, and evaluate fit on the remaining half. Table B.1 shows that the model is able to predict the average selection rate in the test sample quite well.

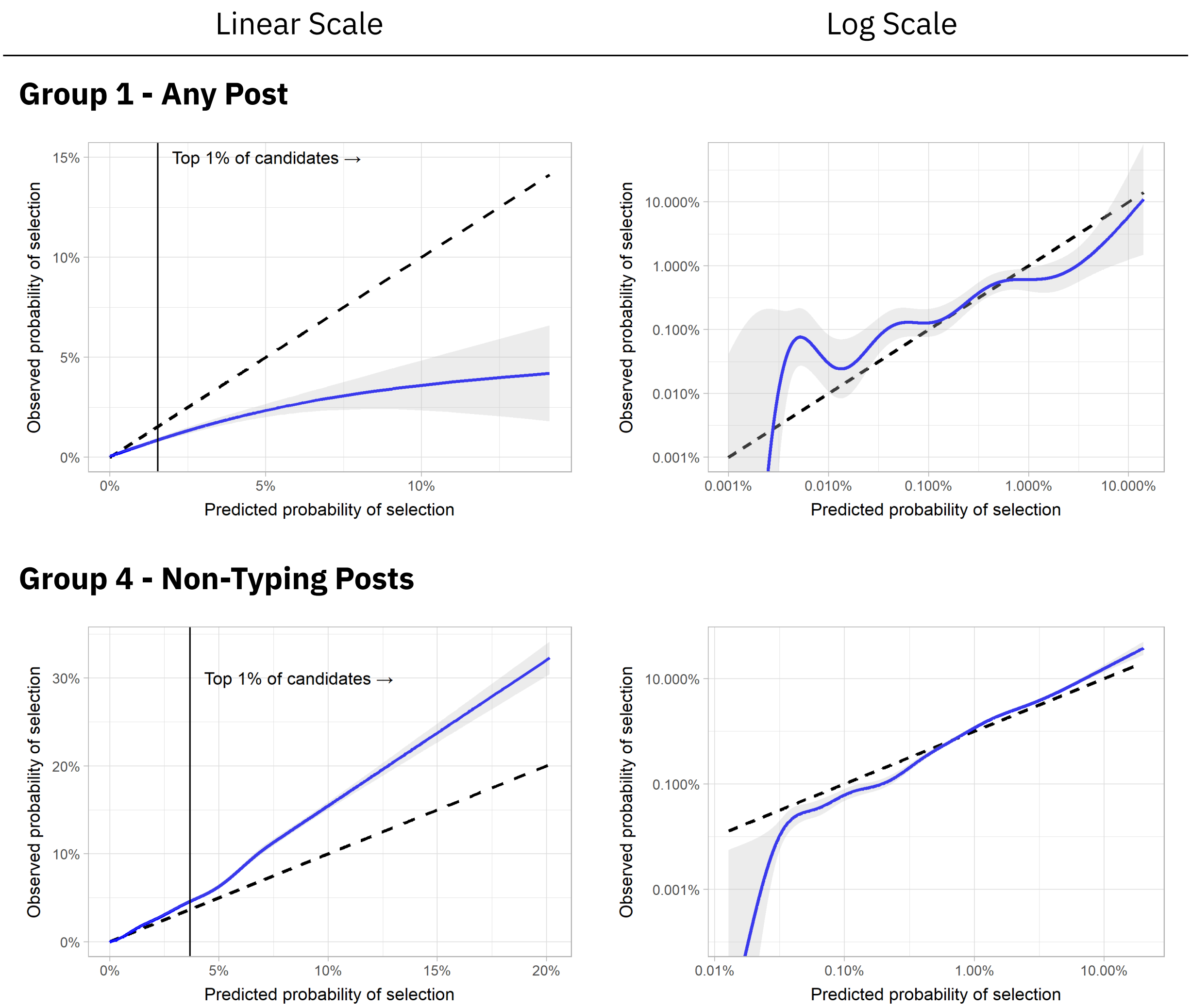
Table B.1: Evaluating Overall Model Fit

Exam Level	Observed Selection Probability	Estimated Selection Probability	% Difference
Group 1	0.1044%	0.1091%	-4.3%
Group 4	0.432%	0.4088%	5.68%

How well does the model fit the data across the distribution? To answer this question, I make use of the following fact. Suppose that the model generates predictions P, and selection outcomes are given by the random variable Y. If the model predictions are unbiased, then $E[Y|P] = P$.

By this logic, I can test the accuracy of the predictions by running a regression of selection outcomes on the predicted probability of selection, and testing whether the regression line is close to the 45 degree line. The results of this analysis is presented in Figure B.1. The two columns show two different views of the same comparison.

Figure B.1: Evaluating model fit across the distribution



Notes: The dashed line is the 45 degree line. The blue line is the average selection rate, conditional on the prediction. The grey band is the 95% confidence interval of the average selection rate conditional on the prediction. If the predictions are accurate, the blue line should land on top of the dashed line. Each column depicts a different view of the same underlying data to account for the large variation on the horizontal axis.

In Group 4, the model matches the observed data well for most of the distribution. However, the fit deteriorates towards the tails. At the low end, candidates appear to have selection rates about 1/100th of what the model predicts. At the top end, candidates have selection rates about 2-3 times what the model predicts. This deviation probably reflects the fact that the included variables limit the amount of heterogeneity the model is able to detect.

In Group 1, the model is also mostly unbiased in the middle of the distribution, though the variance is higher. At the top of the distribution, the model tends to *over-state* the heterogeneity in selection rates by a factor of about 2-3 times.

B.2 Estimating Luck

This section provides additional details on how the luck component of the total score was estimated in Section 3.3.

Estimation Strategy Suppose that candidate i's score T_i on a set of questions can be decomposed as follows:

$$T_i = \mu_i + \epsilon_i \tag{2}$$

where μ_i is the candidate's ability, and ϵ_i is the variation due to luck. Here, μ_i is a scalar quantity, and we assume $\epsilon_i \sim N(0, \sigma^2)$.

Our goal is to estimate σ^2. To do so we, will divide the test scores based on the scores on odd and even questions. The key assumption is that doing so gives us two *uncorrelated* measures of candidate's performance. That is, we assume that:

$$T_i^o = \frac{1}{2}\mu_i + \epsilon_i^o \tag{3}$$

$$T_i^e = \frac{1}{2}\mu_i + \epsilon_i^e \tag{4}$$

$$\epsilon_i = \epsilon_i^o + \epsilon_i^e \qquad Cov(\epsilon_i^o, \epsilon_i^e) = 0 \tag{5}$$

$$\tau^2 \equiv Var[\epsilon_i^o] = Var[\epsilon_i^e] \tag{6}$$

where the superscript o corresponds to variables for odd questions, and the superscript e corresponds to variables for even questions. These assumptions ensure that $T_i = T_i^o + T_i^e$.

Then:

$$T_i^o - T_i^e \sim N(0, 2\tau^2) \tag{7}$$

Since $Var[\epsilon_i] = Var[\epsilon_i^o] + Var[\epsilon_i^e]$, we have $\sigma^2 = 2\tau^2$. We can therefore estimate σ^2 with $Var[T_i^o - T_i^e]$.

The share of total score variation that is explained by luck is given by $\sigma^2 / Var[T_i]$.

B.3 Estimating the Impact of the Information Experiment

This appendix provides additional details on how the impacts of the Information Experiment described in Chapter 4 was estimated.

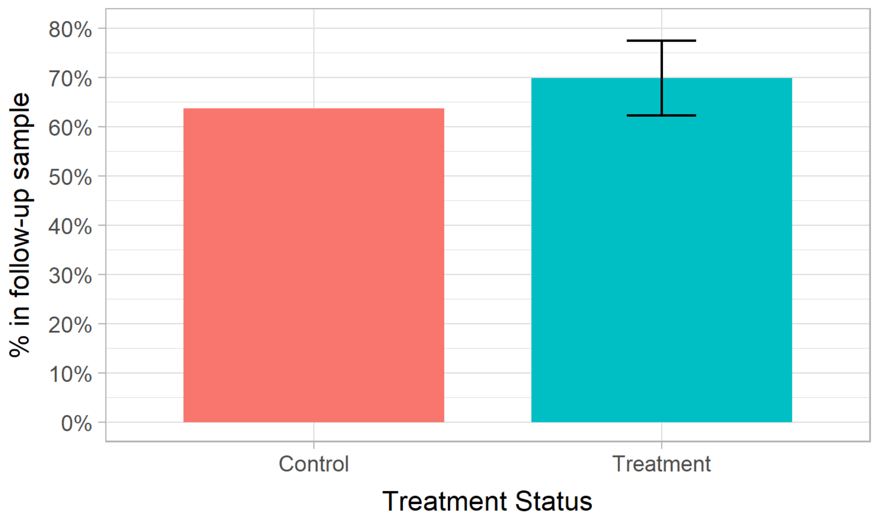
Sample Construction The analysis sample consists of respondents who:

- Were not randomized into the second treatment that was also conducted as part of the Candidate Survey
- Were shown information about the Group 4 exam.[72]
- Had made fewer than two prior attempts in a Group 4 exam
- Completed the Follow-up survey

[72] In case, respondents were applying to both Group 1/2 and Group 4 exams, they were randomly shown information about one of them.

The first three criteria are, by construction, independent of treatment assignment. There could, however, be an endogenous response of treatment to candidates' willingness to participate in the Follow-up survey. The follow-up rate in the Treatment group was about 6.1 p.p. higher (s.e. = 3.9 p.p.), but this difference is not statistically distinguishable from zero (see Figure B.2).

Figure B.2: Treatment assignment is uncorrelated with completing the Follow-up survey



Notes: The figure plots estimates from a specification that includes fixed effects for how the respondents were sourced. The error bar reflects the 95% confidence interval.

Treatment Balance A key assumption for the estimation strategy is that, on average, the only difference between treatment and control groups lies in their exposure to treatment, and not in other pre-determined characteristics.

To test this assumption, we compare treatment and control groups on a range of characteristics that were determined before treatment assignment (see Table B.2). If the treatment assignment is valid, then we should not expect to find many differences.

Table B.2: Treatment balance in the Information Experiment Sample

Variable	Control Mean	Treatment - Control Difference
Age	24.56	0.10
		(0.37)
Female	0.48	-0.0014
		(0.04)
College graduate	0.52	-0.04
		(0.04)
10th Std. Marks	370	14
		(6.3)
Ever worked	0.68	0.02
		(0.04)
# Months preparing	4.20	0.06
		(0.26)
Took prior TNPSC Group exam	0.48	-0.06
		(0.04)
Distance of prior from truth	-0.15	0.04
		(0.02)
Prior belief on cutoff	185	7.51
		(5.29)

Notes: The last column presents the coefficient from a regression of the variable on a treatment indicator and a set of fixed effects for how respondents were sourced. Standard error are reported in parentheses below the coefficient estimate.

The only differences that are statistically significant are in candidates 10th standard marks, and in the distance between the prior and the truth. We control for both of these variables in the main specification.

Regression Specification We estimate the impact of the Information Experiment using one of the following regression specifications:

$$y_i = \alpha_{s(i)} + \beta \text{Treat}_i + \Gamma \mathbf{X}' + \epsilon_i \tag{8}$$

$$y_i = \alpha_{s(i)} + \beta_1 \text{Treat}_i + \beta_2 \text{Pessimist}_i + \beta_3 \text{Treat}_i \times \text{Pessimist}_i + \Gamma \mathbf{X}' + \epsilon_i \tag{9}$$

The coefficient $\alpha_{s(i)}$ captures a fixed effect for how the respondent was sourced, i.e. whether they came through Facebook or some other means, and if through Facebook which ad they clicked.

The vector $\mathbf{X}$ includes all the controls chosen in this regression, including:

- A cubic polynomial in the distance between the prior and the true value
- The candidates' age
- Candidates' prior beliefs on the cutoff score
- Candidates' 10th board exam marks

These control variables were selected from a longer list of potential controls using the method of causal forests proposed by Wager and Athey (2018), and implemented in the R package `grf` (Friedberg et al., 2020).

Observations are weighted so that the estimates are more representative of the population of candidates as a whole.

B.4 Estimating the Item Response Theory Model

This section provides additional details on how the IRT model was estimated in Section 5.3.

The 3PL Model I fit what is known as a 3 parameter logistic curve to the data. For each person i, the model assumes that the probability that they answer question k correctly is given by the following equation:

$$Pr(y_{ik} = 1 \mid \theta_i, a_k, d_k, g_k) = g_k + \frac{1 - g_k}{1 + \exp\left(-a_k(\theta_i - d_k)\right)} \quad (10)$$

where

- a_k is the item discrimination parameter
- d_k is the item difficulty parameter
- g_k is the item guessing parameter
- and θ_k is the candidate's ability

If there are N individuals and K test items, then there are a total of $N + 3K$ parameters to estimate.

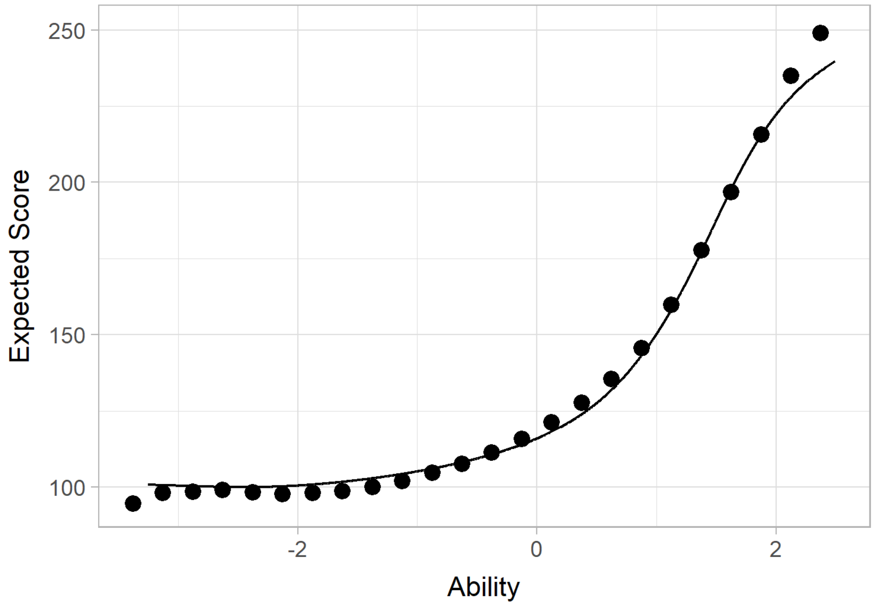
The estimates of θ_i may not rank candidates in the same way as ranking them based on the numbered answered correctly. The reason is that not all questions have the same *discrimination*. The estimate of θ_i will place more weight on the items with a higher a_k since these are more informative.

Estimation The model is estimated via Maximum Likelihood. Since the Likelihood function depends on unknown state parameters (namely the θ_i), the maximum likelihood estimates do not have a closed form solution.

I estimate the model using an Expectation Maximization algorithm, as implemented in the R package `mirt` (Chalmers, 2012). The estimation procedure simultaneously generates a non-parametric estimate the latent distribution of θ_i along with estimates of the item parameters.

Model Fit Figure B.3 compares the fit of the model with the observed data. In general, the model fits the data quite well, though it does under-predict candidates' scores at the top of the distribution. This suggests that more work may be needed to find better parameterizations or estimation strategies in order to account for this deviation.

Figure B.3: Comparing the IRT Model Fit with the Observed Data



Notes: The solid line is generated by the model. The points come from the data. Candidates with scores below the minimum qualifying threshold have been dropped from the estimation procedure. Sample is restricted to candidates taking the Tamil version of the test.

Simulation Exercise Once we have a parametric model for the probability that a candidate answers a question correctly, we can use it to simulate how the score distribution would change when the model parameters change.

This is the approach I use in Section 5.3 to simulate how the distribution of scores at the top of the distribution would change if question quality improved.

This simulation proceeded in two steps. First, I generated 200 simulated questions by randomly sampling the following parameter values (independently of each other):

- $a_k \sim$ Uniform$(1.2, 2.5)$
- $d_k \sim$ Uniform$(1, 2.5)$
- $g_k \sim$ Uniform$(0, 0.1)$

Second, using the already estimated θ_i and the simulated IRT parameters, I generated a simulated response pattern. To do so, I drew a Bernoulli random variable for each candidate and each question where the probability of a correct answer was given by plugging in the parameters into Equation 10.