Building the Evidence Base on Urban Transport

Interim Results from ADB-Supported Impact Evaluations in Lahore, Pakistan

Kate Vyborny

Duke University

April 2018

Motivation

- Rapid urban growth and increase in private vehicle use: **environmental concern**
- Poor have limited access to cars or public transport; public transport often low quality (speed, reliability, safety): **equity concern**
- In Pakistan, women even more constrained because of social norms and safety: **gender equity concern**
- Major **mass transit expansion** in progress across Pakistan

Two impact evaluations on urban transport in Lahore, Pakistan, conducted by academic research team and supported by ADB:

- **Mass transit line (quasi-experiment)**
- **Pick-and-drop to work, with gender focus (randomized experiment)**

Work in progress

- Some results presented in public documents - for circulation
- Other results are work in progress and subject to change
- Comments welcome - katherine.vyborny@duke.edu

Impact evaluation 1: Mass transit

- Joint work with Hadia Majid (Lahore University of Management Sciences) and Ammar Malik (Urban Institute)
- Working paper available online: "Infrastructure investments and public transport use: Evidence from Lahore, Pakistan"

Background

- Mass transit project highly controversial: need for **independent evidence**
- Completed study focuses on switch to **sustainable commuting**, i.e. commuting by public instead of private vehicle
- Ongoing study explores impact on **other economic outcomes**: markets for labor (workers, jobseekers, firms, small businesses) and land (real estate prices, density)

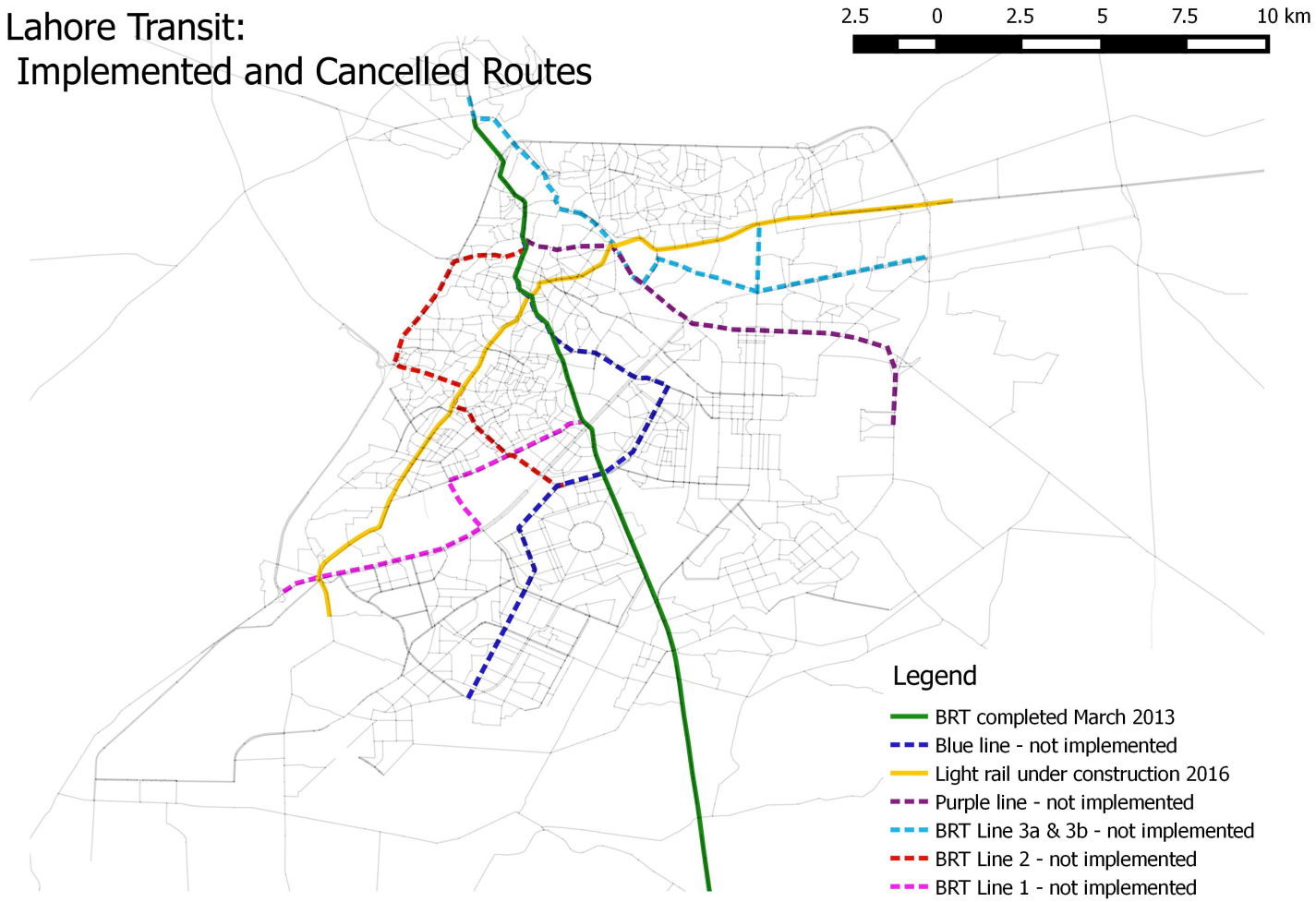
Literature

- Literature on **U.S. mass transit investments** (Baum-Snow and Kahn 2000 JPE, 2005, Winston and Maheshri 2007 JUE):
  - Are built in **lower-income areas**
  - Have **limited impact** on **public transport ridership**; most users switch from buses
  - Do **not justify** capital costs of rail
- More limited evidence from developing world, where:
  - **Larger base** of potential public transport users
  - **Congestion higher** so greater time savings relative to travel in car or bus

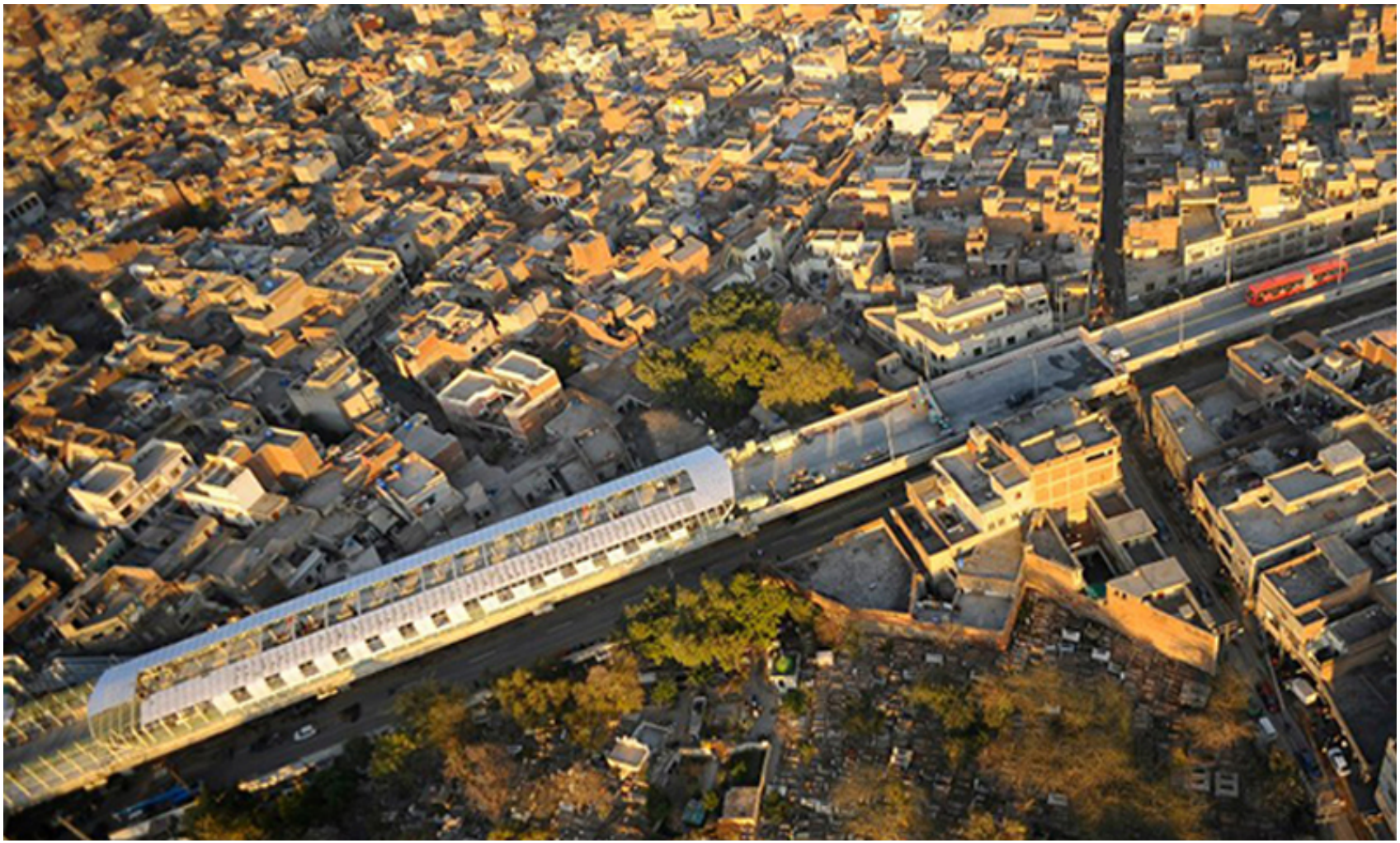
Our contribution

- Challenge of isolating **cause and effect**: government targets transport to areas that might be different for other reasons
- Isolate effects of metrobus with **quasi-experiment** using **similar areas** served by **planned lines** as comparison group
- Detailed **microdata on comparable areas**
- Complement with rich **descriptive data** from household and rider surveys

2007 Lahore mass transit plan

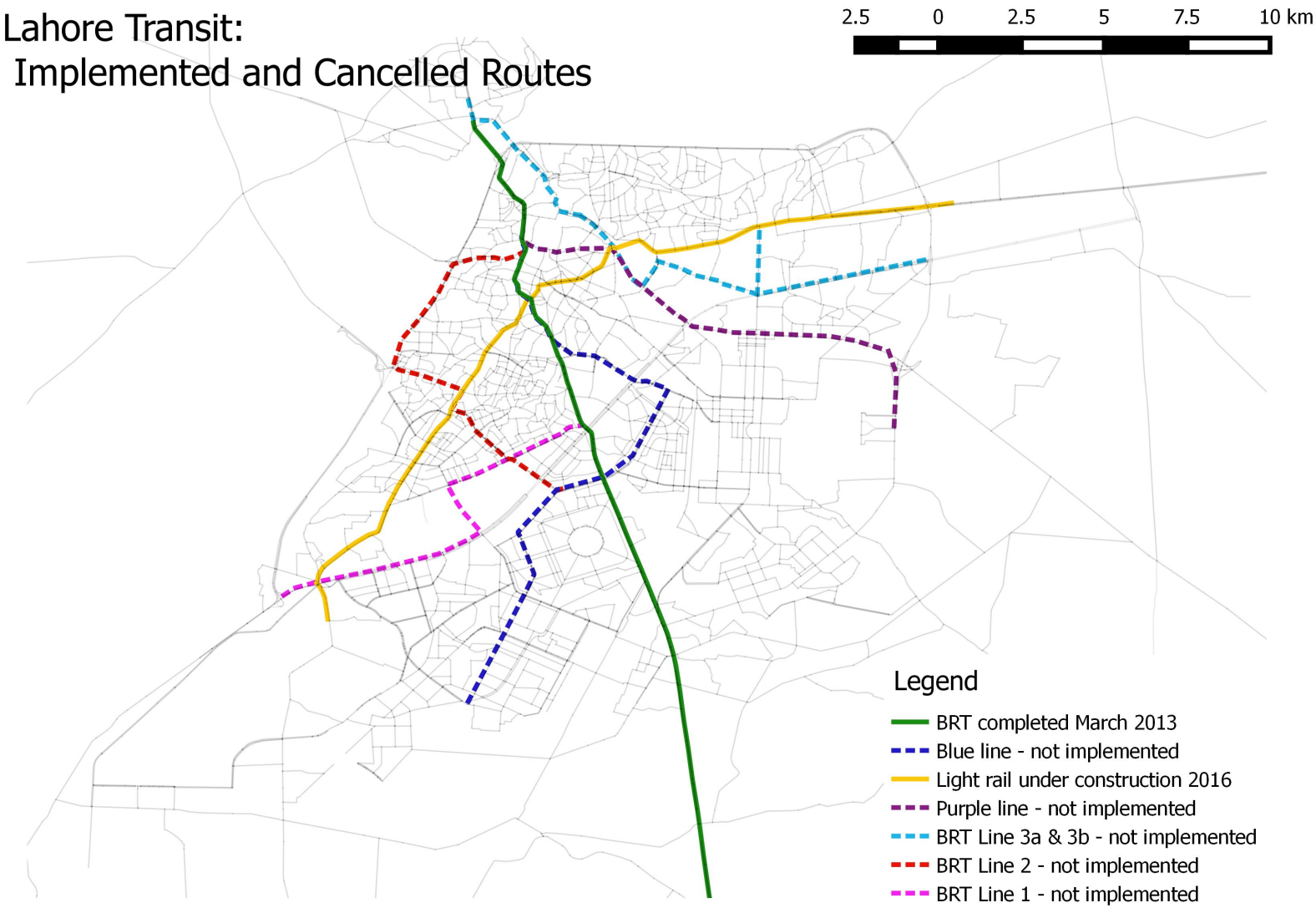


2013 metrobus (green line) crosses entire city N-S

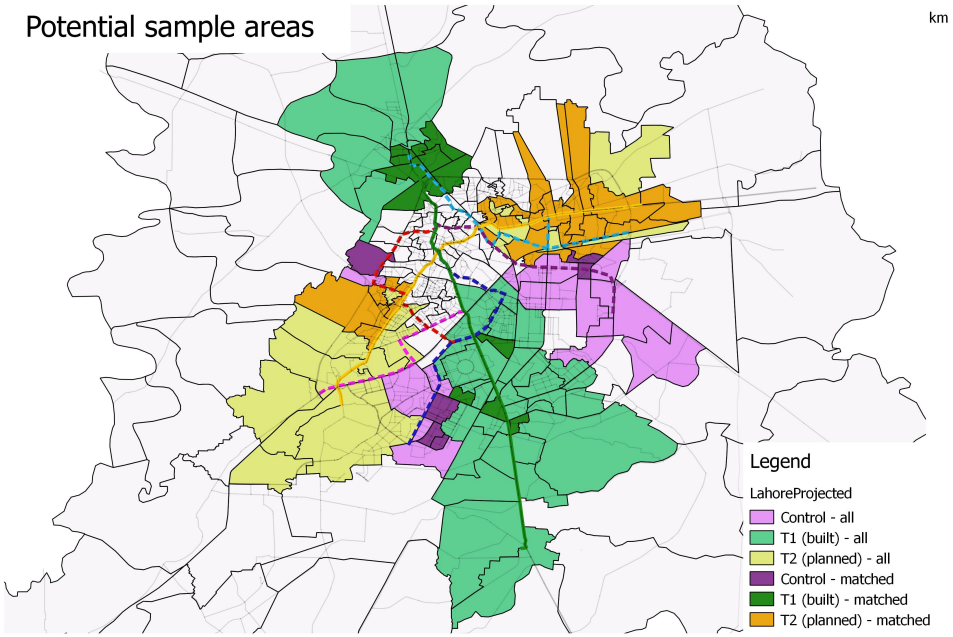


2013 metrobus (green line) crosses entire city N-S

2016-present: orange line under construction

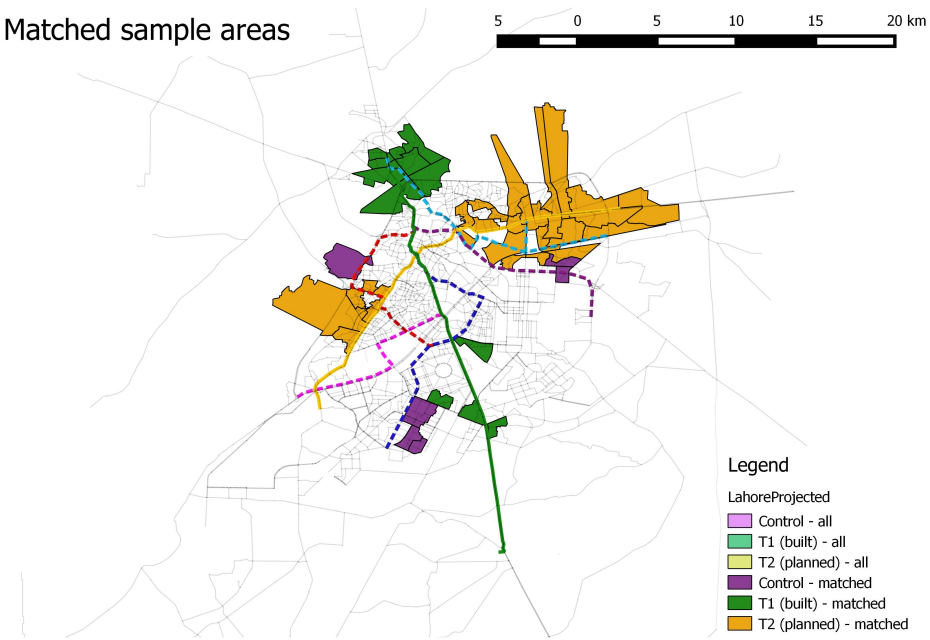

Lahore Transit:
Implemented and Cancelled Routes
2.5 0 2.5 5 7.5 10 km
Legend
BRT completed March 2013
Blue line - not implemented
Light rail under construction 2016
Purple line - not implemented
BRT Line 3a & 3b - not implemented
BRT Line 2 - not implemented
BRT Line 1 - not implemented

Quasi-experiment plus matching



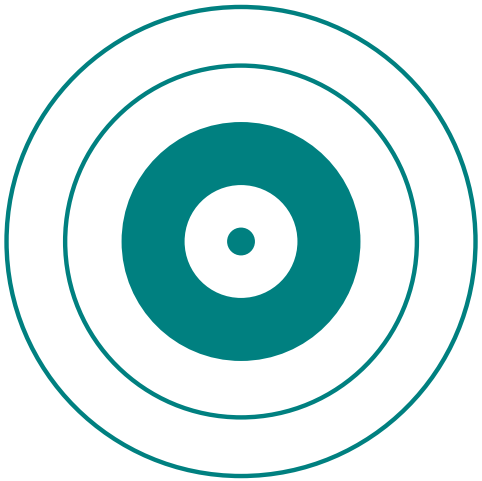
- **Planned lines** are used as **comparison group**
- **Avoid spillovers**: C zones at min distance from T1 / T2 stops

Matched sub-sample



- Select zones **similar on observables** at baseline
  - 2010 data (labor, income, commute, demographics), 1998 census data (pre-trends), characteristics of surrounding zones
    ▸ Details
- **Well-balanced** after matching ▸ Table

"Donut" fixed effects



Built (T1) station

Planned (T2) station

Cancelled (C) station

Identifying assumption for causal effect

- Areas at **same distance to a planned stop and a built stop** with **same characteristics at baseline** do not differ on unobservables that affect commuting
- Robustness to recall-based panel: **parallel trends** within this matched quasi-experimental group

Data

- Causal effect: **Matched comparison groups**:
  - 12,000 **HHs**
  - 500 **real estate agents** / community respondents
- Descriptive analysis: Representative survey of 2,400 **Metrobus riders**

Empirical specification

- X of interest: **Public transport access** reported by community respondent: "How long does it take to get from here to Kalma Chowk using only walking, bus, wagon or metrobus?"
- Use **distance to built stop** as an **instrument for public transport accessibility to central Lahore**

Empirical specification

$$ACCESS_{gt} = \pi_1 + \pi_2 D1 + \pi_3 D1POST_t + \pi_4 D + \pi_5 DPOST_t + \eta Z_g \cdot POST_t + \upsilon_{gt}$$

$$Y_{igt} = \beta_1 + \beta_2 \widehat{ACCESS}_{gt} + \beta_3 POST_t + \pi_4 D + \pi_5 DPOST_t + \gamma X_i \cdot POST_t + \zeta Z_g \cdot POST_t + \epsilon_{igt}$$

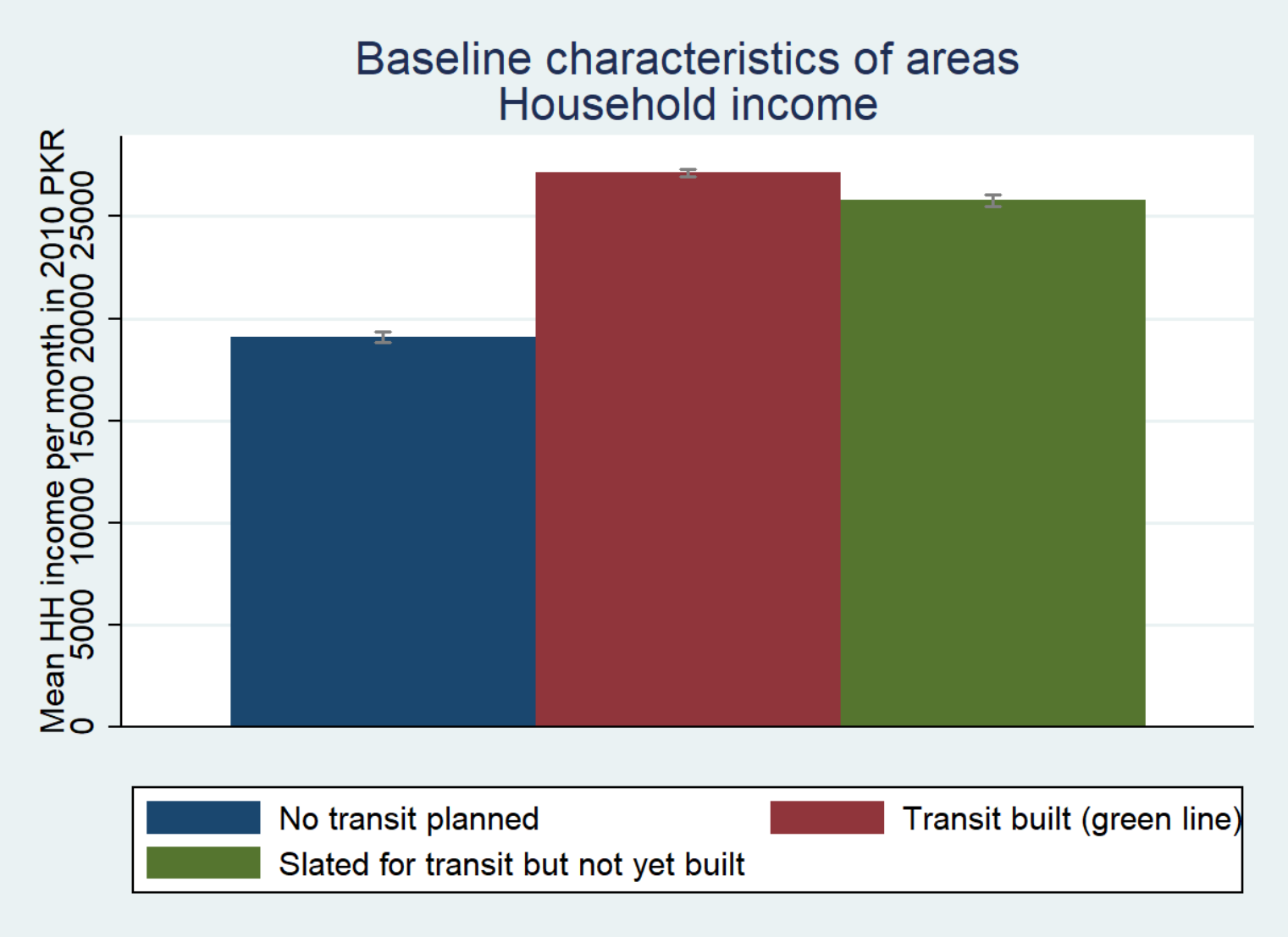
- *ACCESS*: fastest available route on public transport from area g to Kalma Chowk
- $D1$ distance to built stop (T1)
- D distance to any planned stop (T1, T2, C)
- X_i individual control variables
- Z_g group control variables including the "donut rings"
- Cluster ϵ at the zone level

Results: public transport access

On average, each km closer to a transit stop:

- decreases travel time to center on public transport by 3.5 min
- decreases travel fare to center on public transport by 3 PKR

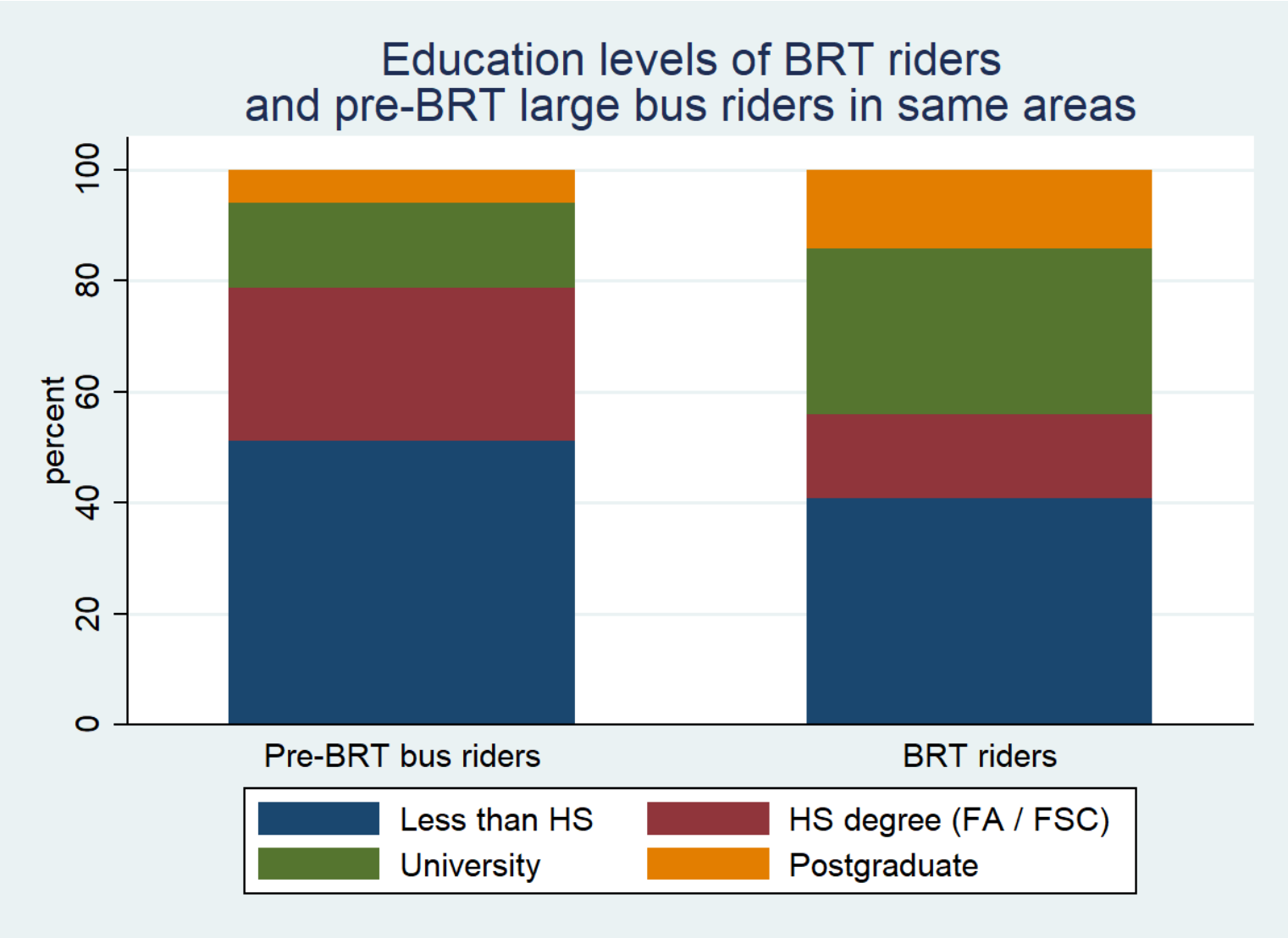
Transit targeted at higher SES areas



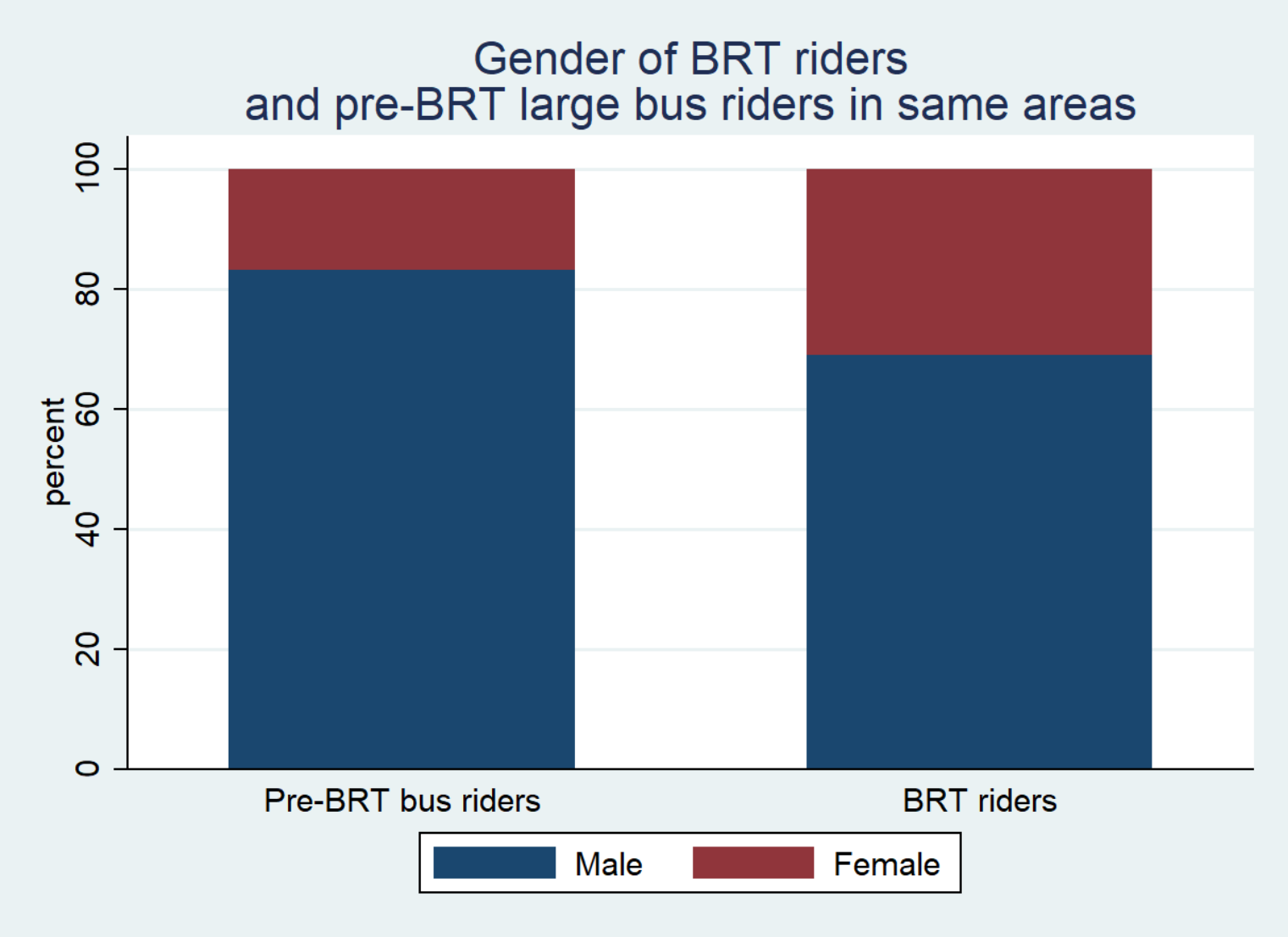
Transit targeted at higher SES areas

- Contrast to US where transit targeted at low income areas
- Likely related to proximity to major roads where transit feasible

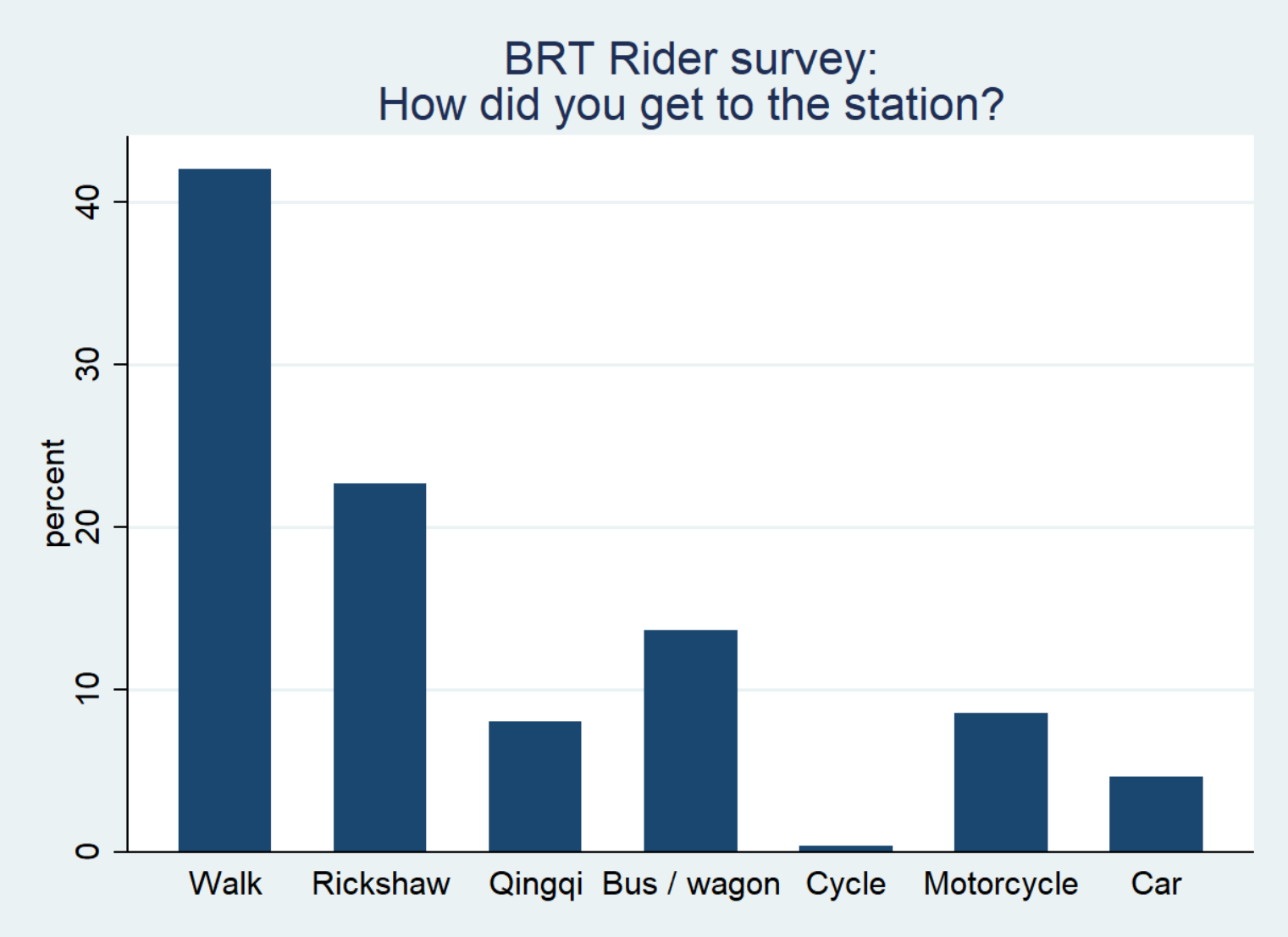
Transit draws higher SES riders than buses



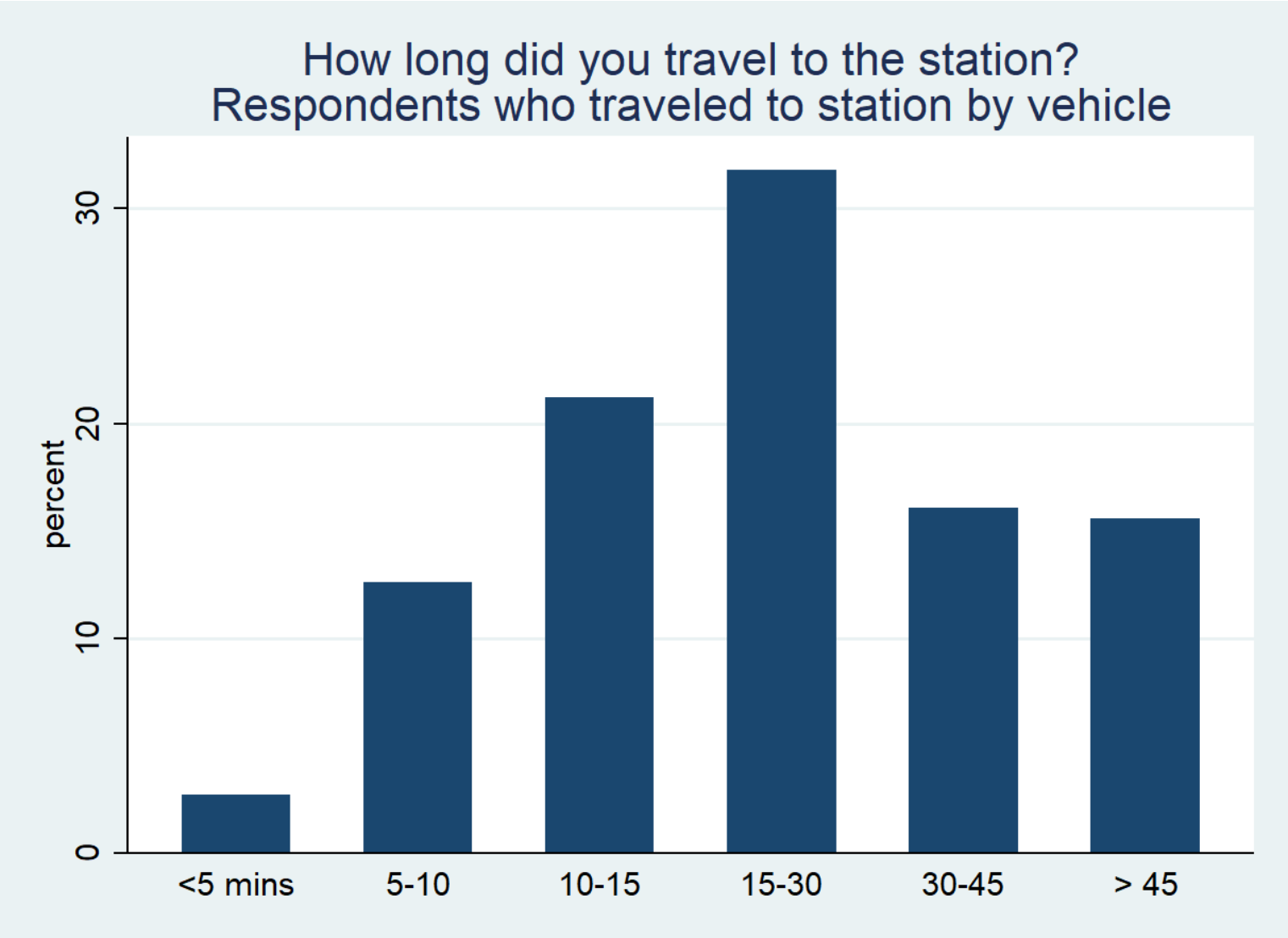
Transit draws more women riders than buses



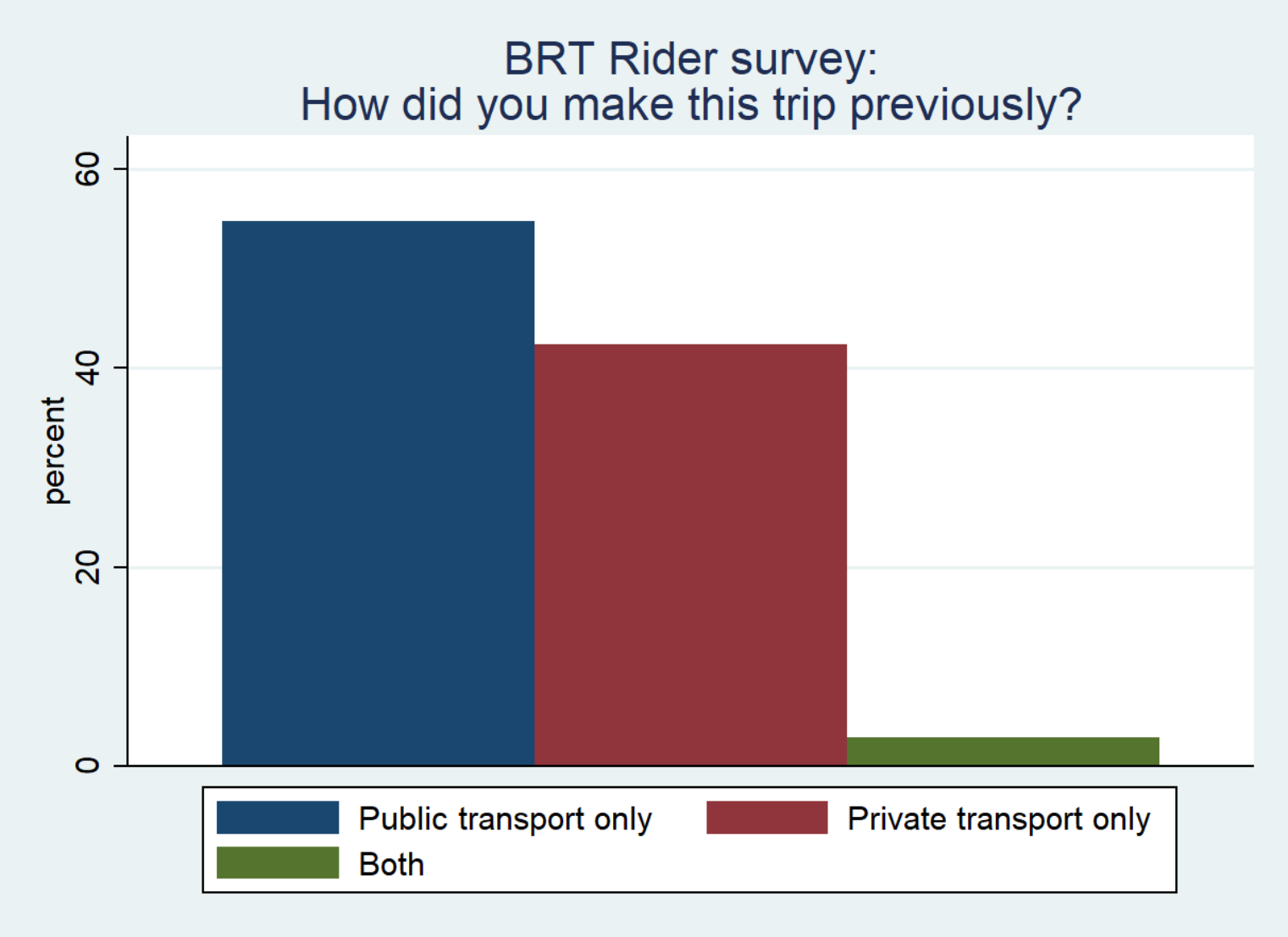
Transit draws riders from far away



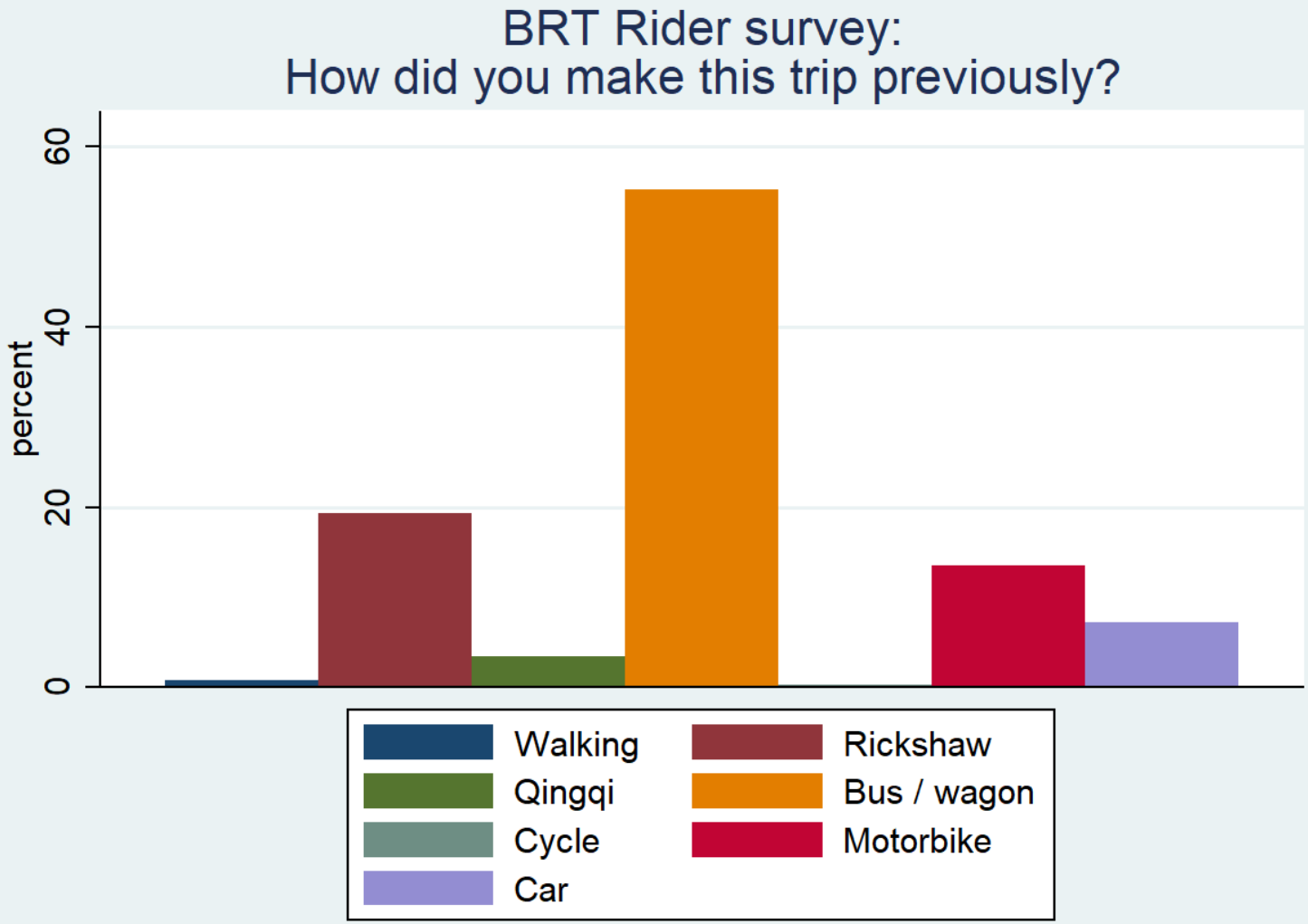
Transit draws riders from far away



Transit riders report switching from private transport



Transit riders report switching from private transport

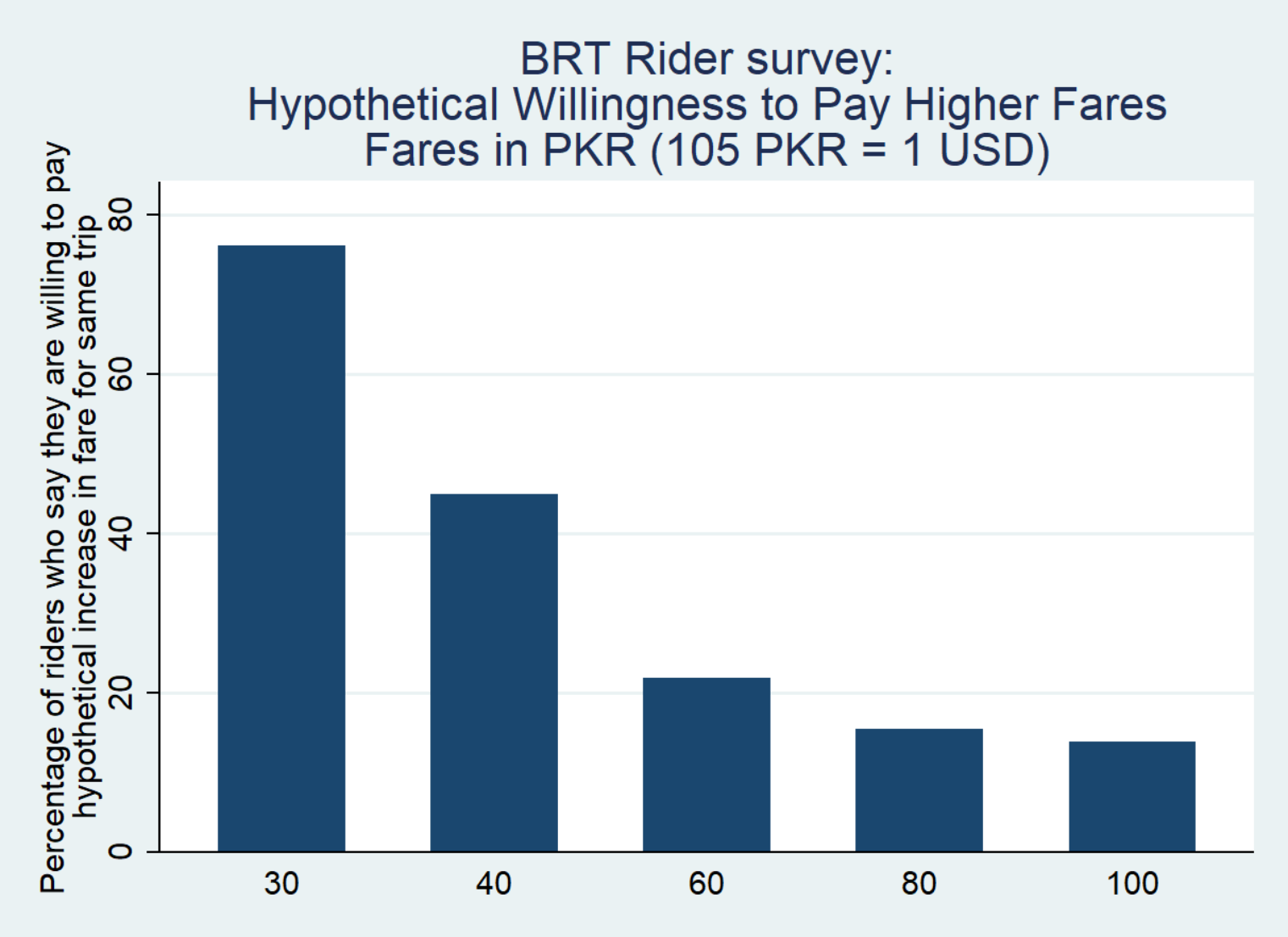


Results: switching to public transport

Causal results (regression estimates):

- Transit caused **significant switch to public transport**
- For areas with faster access to central Lahore via Metrobus, every **10 minute increase in public transport access to center** increases the number using public transport by 1.3 percentage points, or about **20% increase** over base of 7%
- Estimate **35,000 switchers overall**
- Most switch from **motorbikes**
- Estimate about **6,000 tons CO2 averted per year** from switching

Riders willing to pay much higher fares



Conclusions on targeting and public transport use

- Metrobus attracted **more high SES riders and more women**
- Mass transit caused **significant switch from private to public transport**
- Riders are willing to pay substantially more; **subsidy could be better targeted**

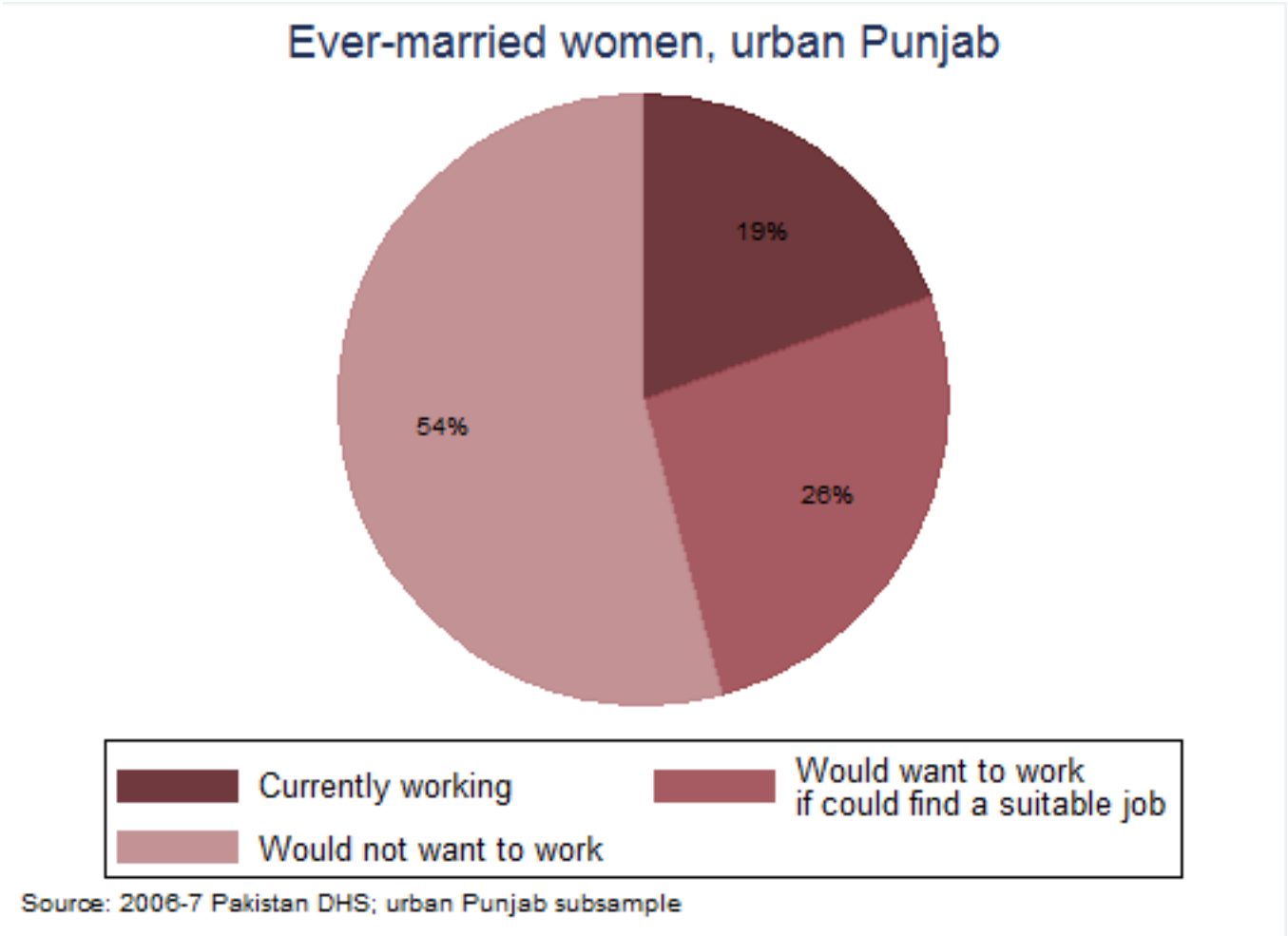
Economic activity: **work in progress**

- Increase in **rents** in high density areas
- Increase in **urban density** near stations
- **No change** in **overall workforce participation**
- Increase in activity of small **local businesses near stations**
- Ongoing work: using firms survey and jobs matching platform to study firm and worker impacts in more detail: data on **firm-worker match quality** to separate **overall gains vs. change in spatial pattern of activity**

Impact evaluation 2: Transport randomization

- Joint work with Erica Field (Duke University)
- Policy brief from first phase available online: "Overcoming Barriers to Women's Mobility: Improving Women's Access to Public Transport in Pakistan"

Few women work, but many women say they **want to work**

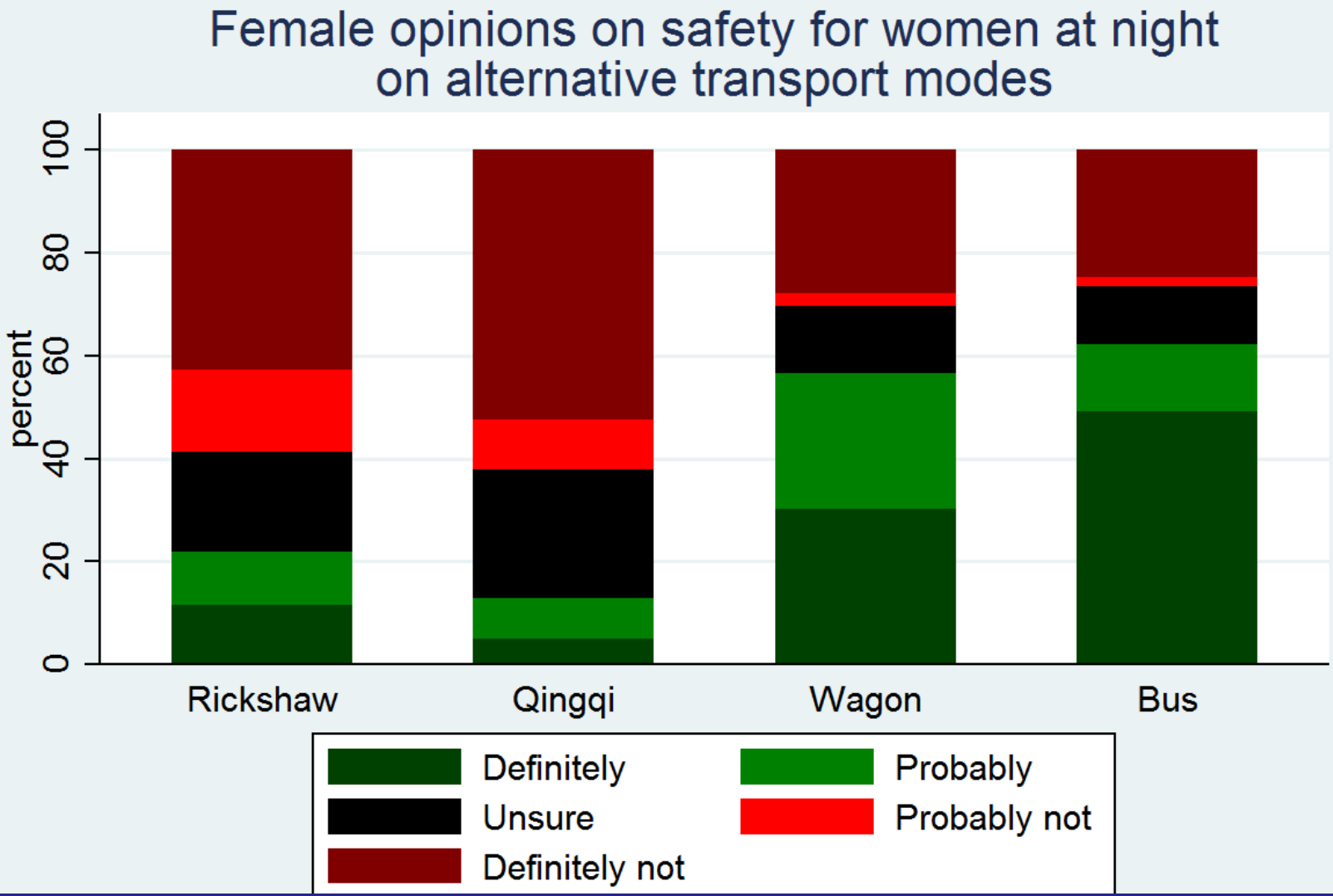


Source: 2006-7 Pakistan DHS; urban Punjab subsample

Getting to work is one of the major challenges

> "Sometimes men follow me to the bus stop - that's just how things are!"
>
> - Worker, garment factory, Kot Lakhpat, Lahore

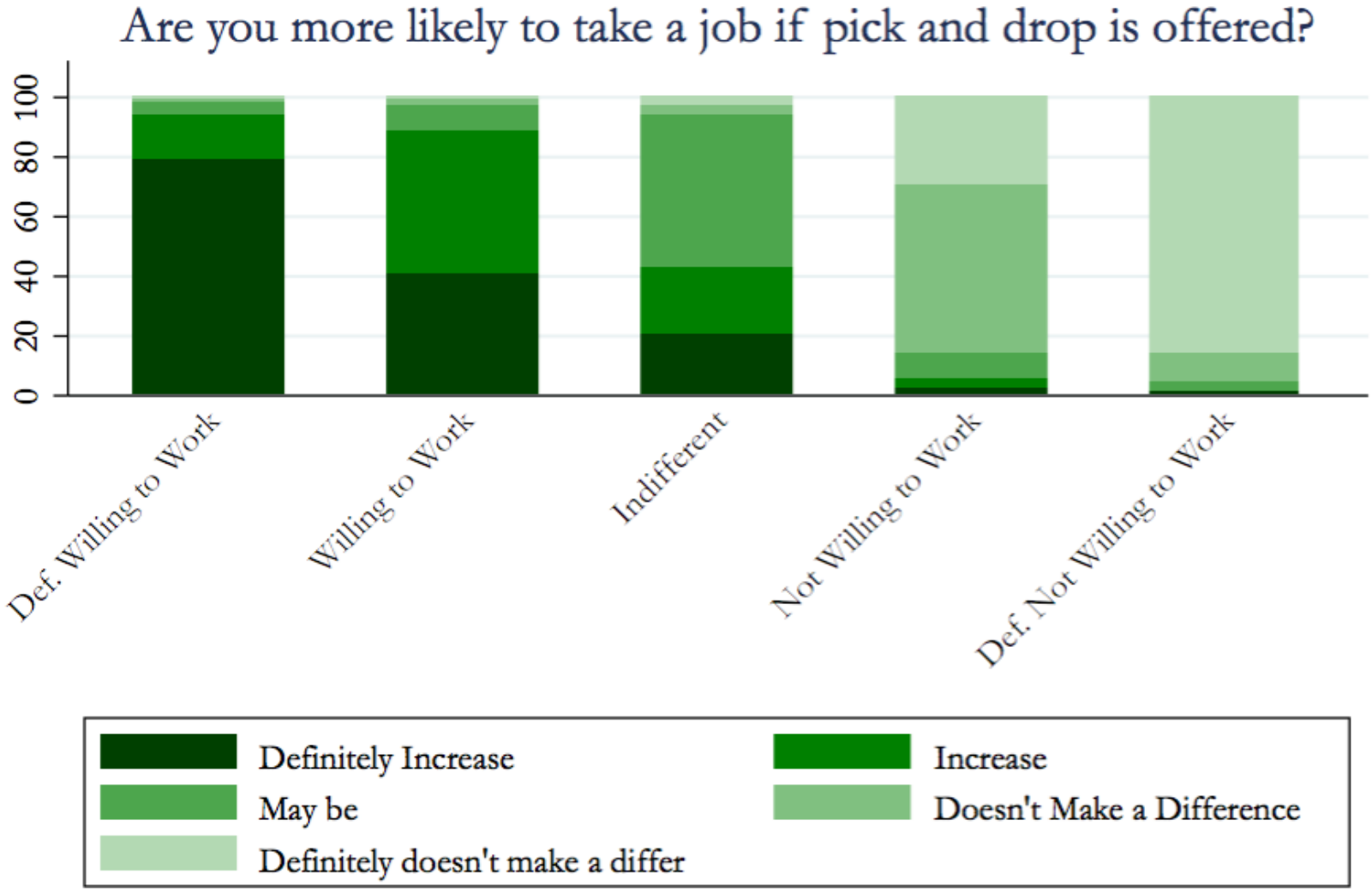
Getting to work is one of the major challenges



Women say transport affects their choice of jobs

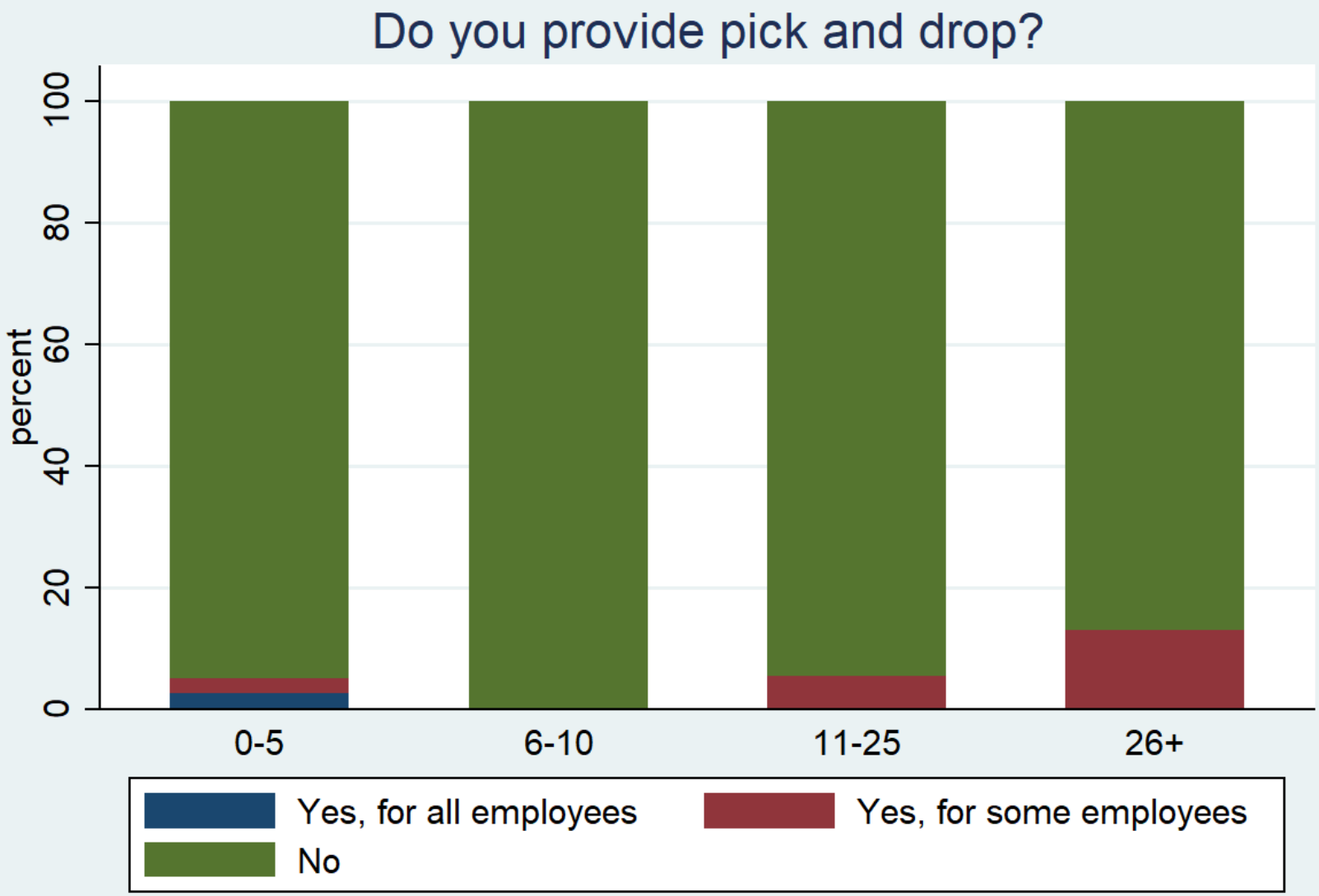
> "The factory I worked at before did not offer me conveyance, so coming and going from home to work was very problematic for me. That's why I changed my job to this one."
>
> - Worker, garment factory, Kot Lakhpat, Lahore

Women say transport affects their choice of jobs



Def. Willing to Work = 2822, Willing to Work = 3266, Indifferent = 3532,

Only a few larger employers provide transport



Policy brief on gender and mobility in Pakistan - from background work to IE

Policy brief ("Overcoming Barriers to Women's Mobility" - CDPR):

- Extend **high-quality public transport** with less crowding
- **Make streets safer for women**: police, night lights
- Continue and evaluate efforts to **normalize women in public space** ("women on wheels")
- Train and monitor transport staff on **harassment**

Randomized Controlled Trial

- **Baseline surveys** of men, women and employers
- Designate routes:
  - Group 1: **female-only** pick and drop services
  - Group 2: **mixed-gender** pick and drop services
  - Group 3: **"Control" routes** for comparison
- Provide **pick-and-drop for one year**

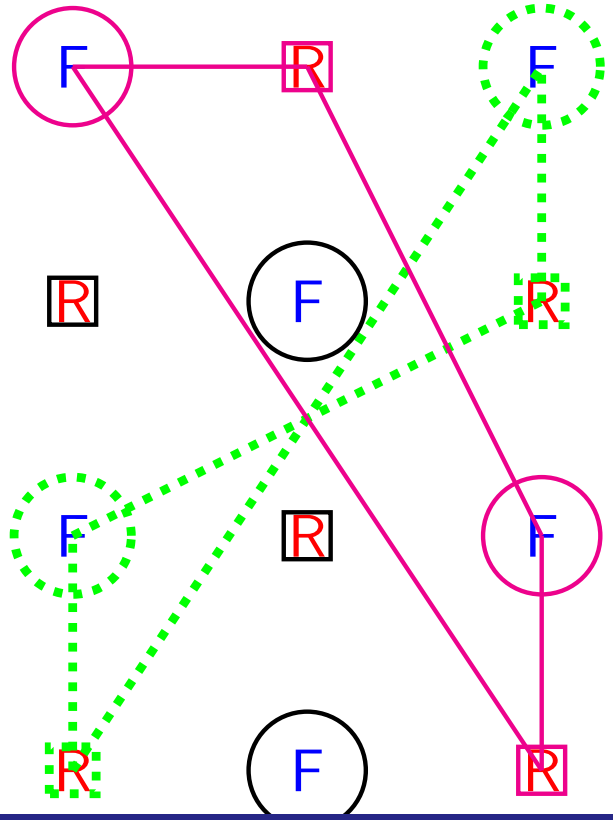
Phase II: Randomized Controlled Trial

- **Follow-up survey**: did service improve outcomes for **male, female jobseekers and employers**?
- Test and evaluate **one approach**: if successful, could be scaled up and form part of the solution
- Quantify potential economic benefits of **all policies** that can improve mobility (not just pick / drop)

Design

Treated residential areas: Women's-only transport

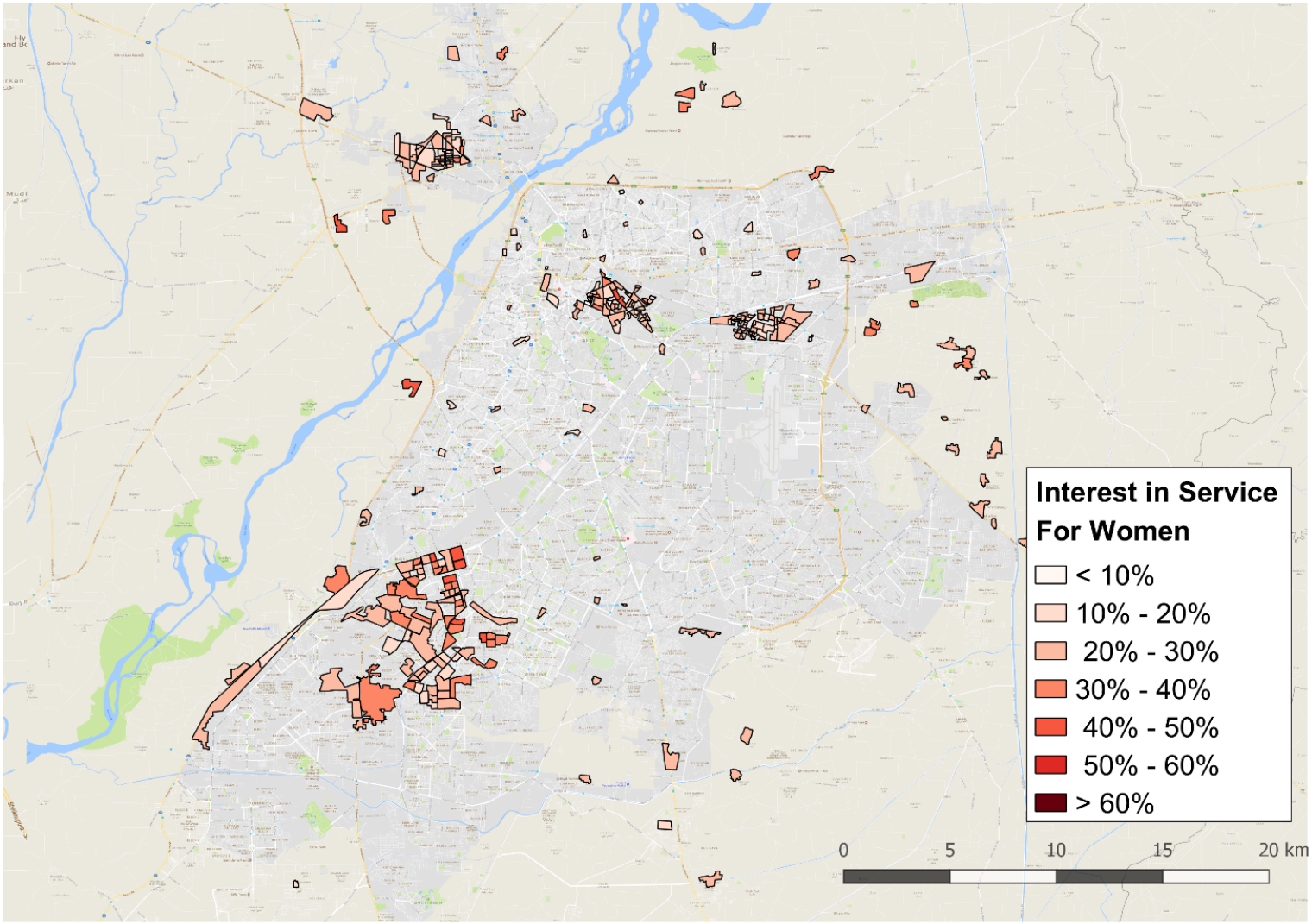
Treated residential areas: Mixed-gender transport



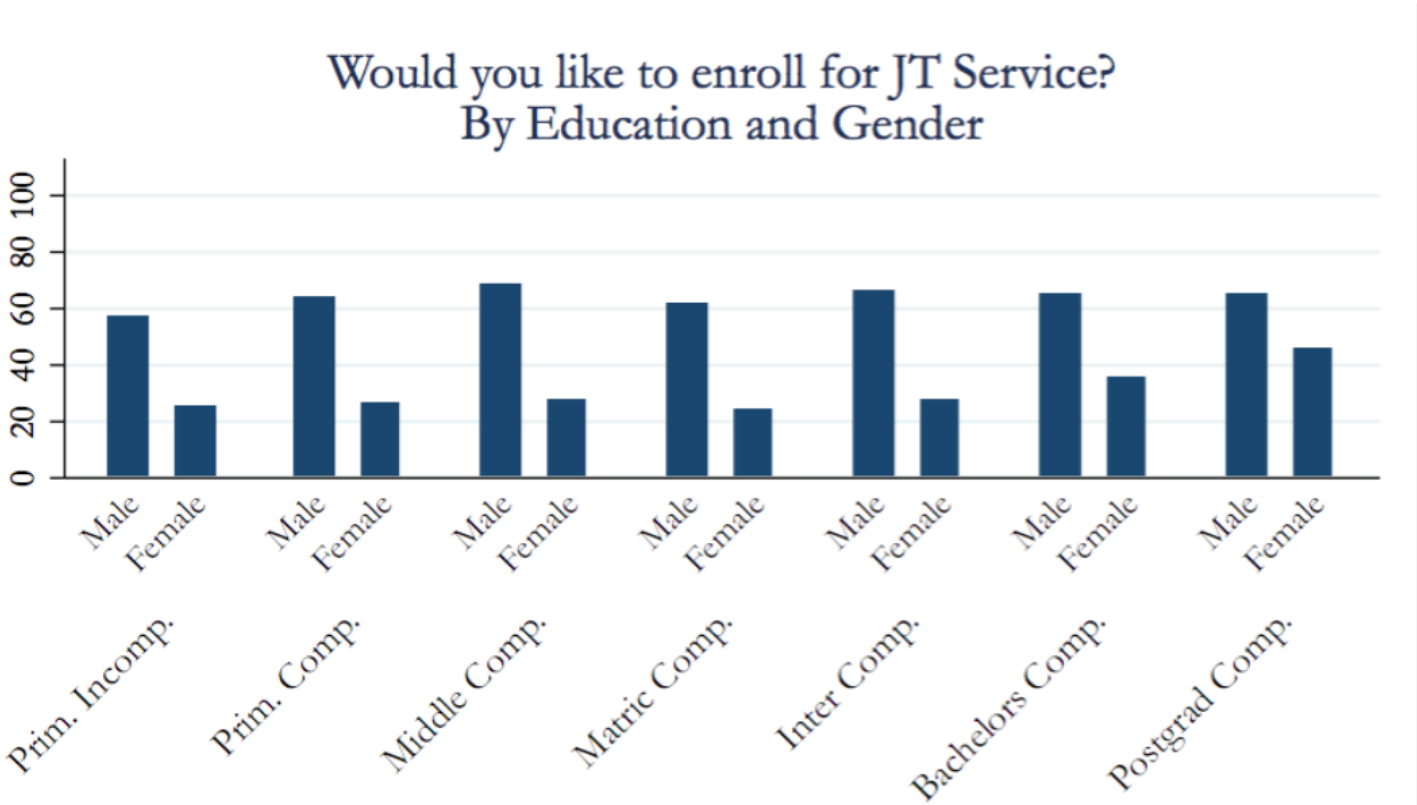
Measurement strategy

- Measure whether men / women **apply to and get better jobs further from home**
- Measure whether firms have access to a **better pool of employees**
- Research team developed **"Job Talash"** (Job Search) service
- Offered to all HHs and employers at baseline

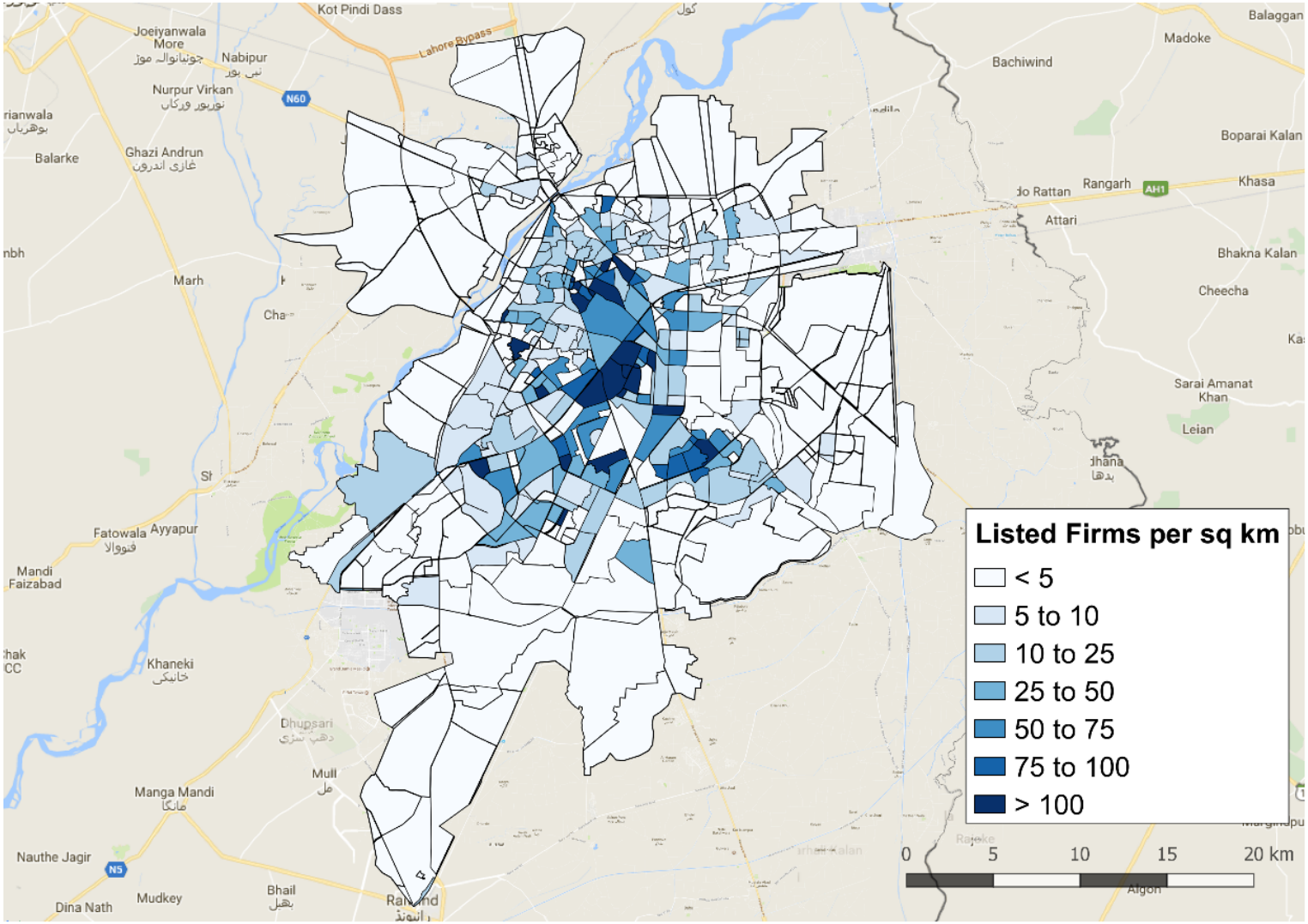
Strong interest in Job Talash - urban and periurban



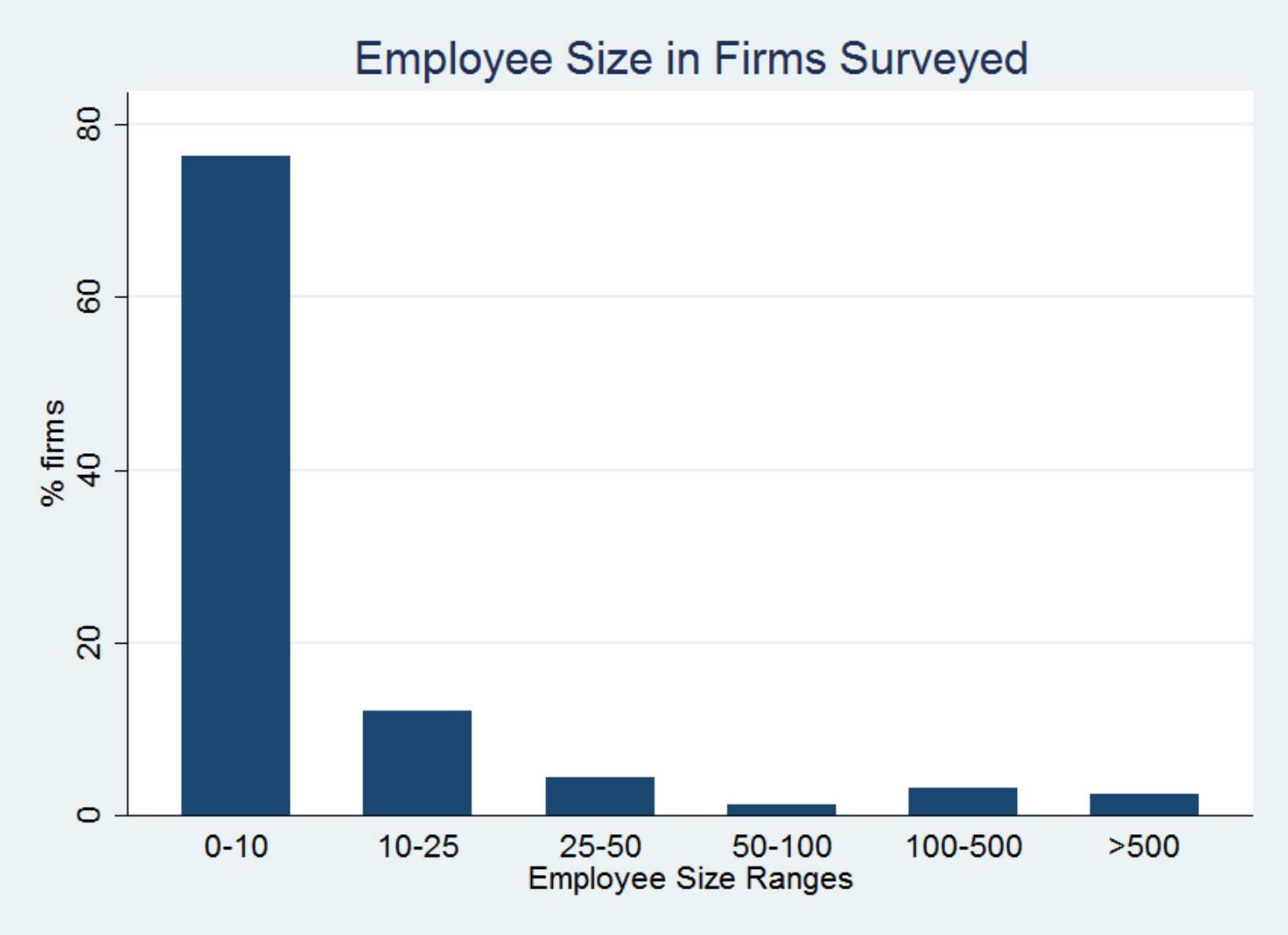
Interest across gender and ed; increasing in ed for women



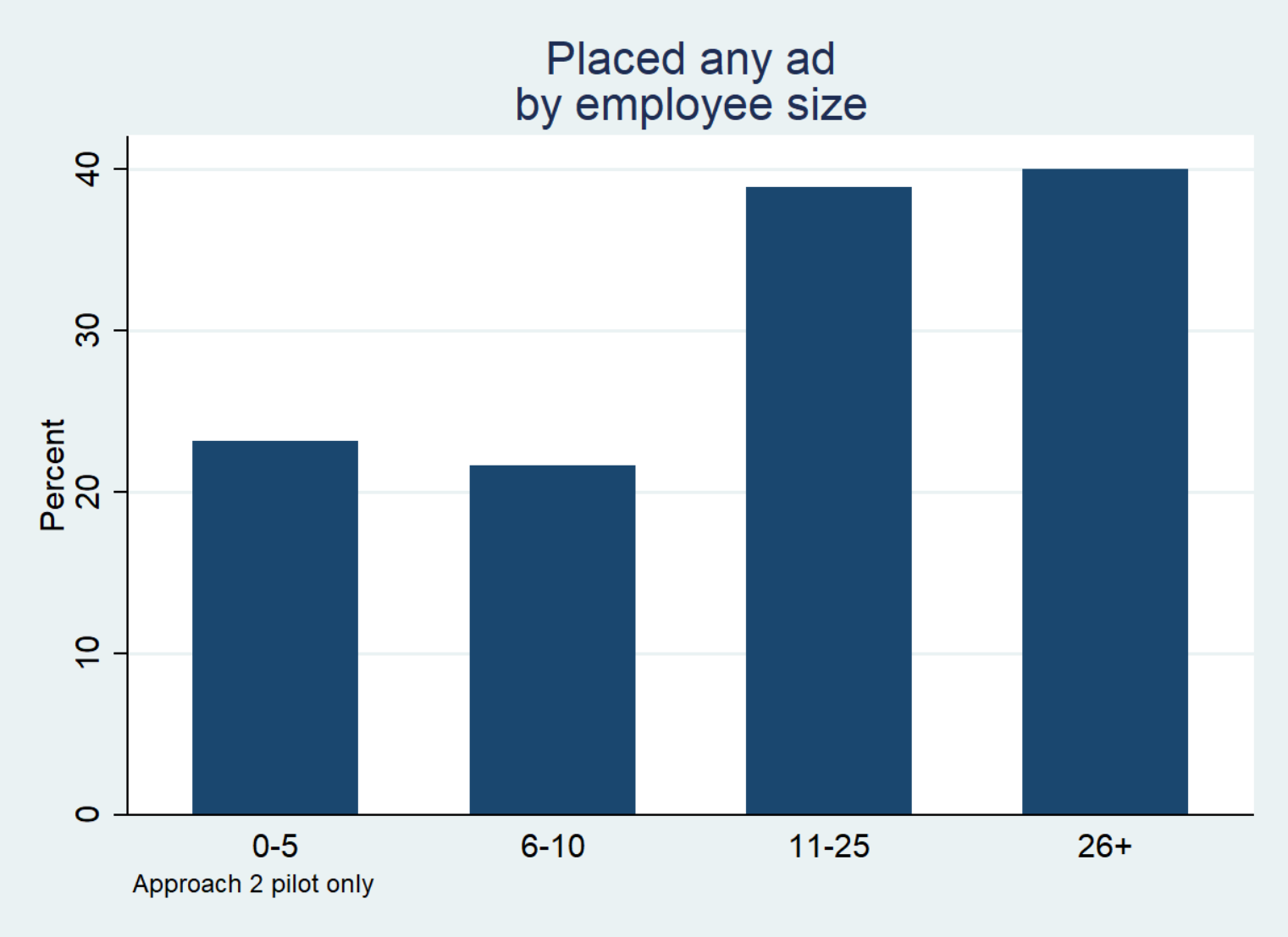
Employers are spread across metro area



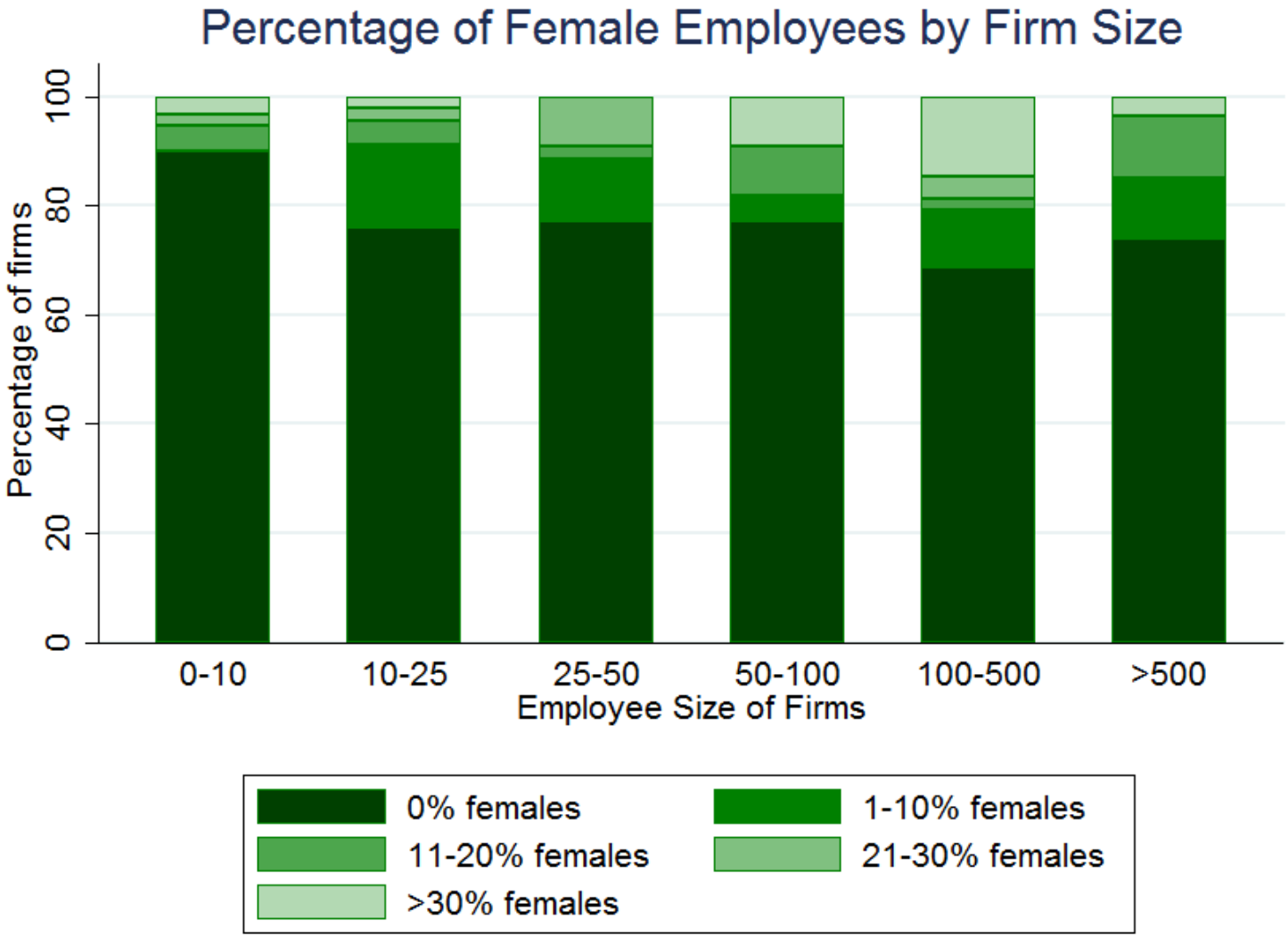
Field survey representative of all employers



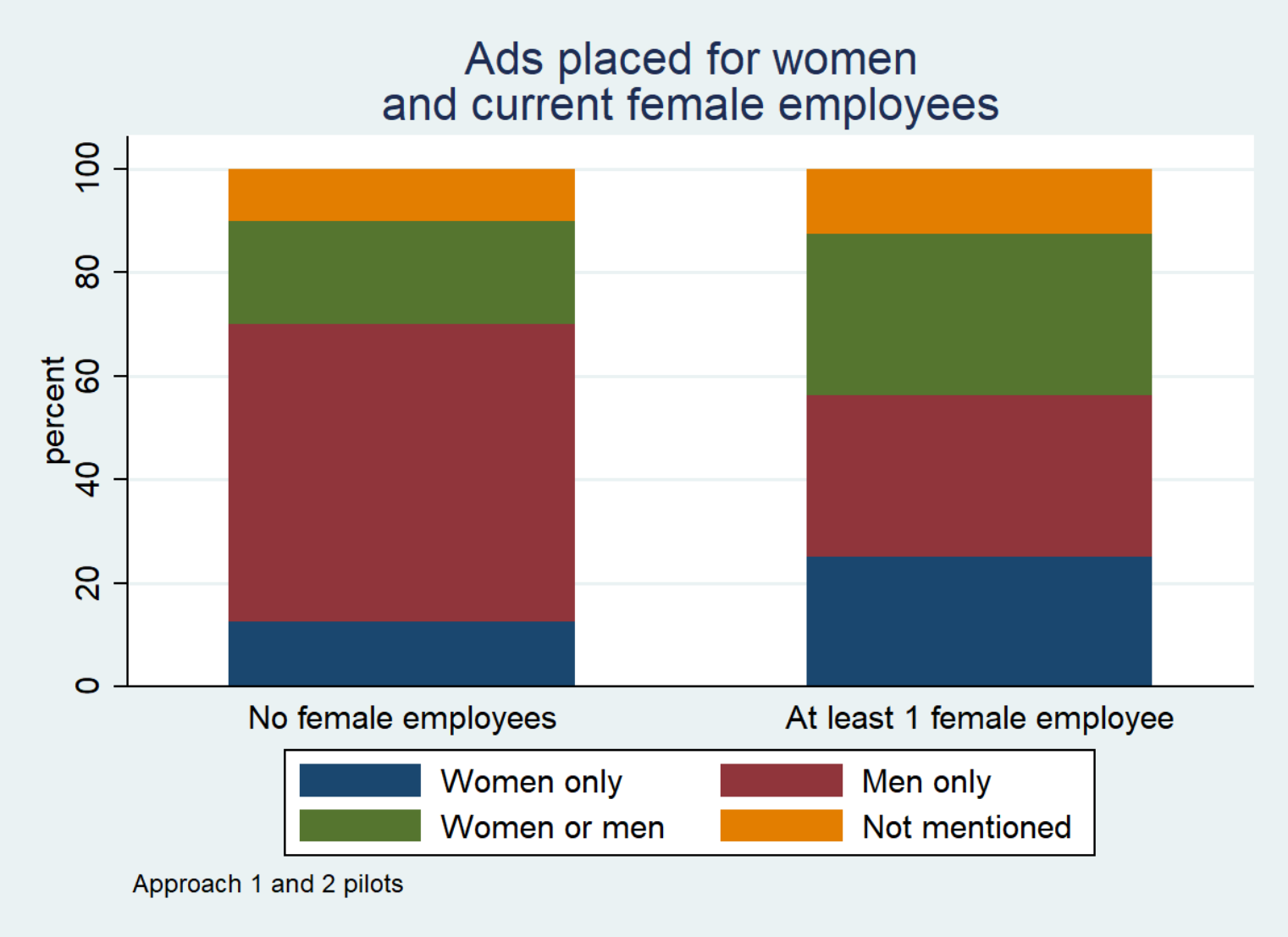
Firms of all sizes want Job Talash to send candidates



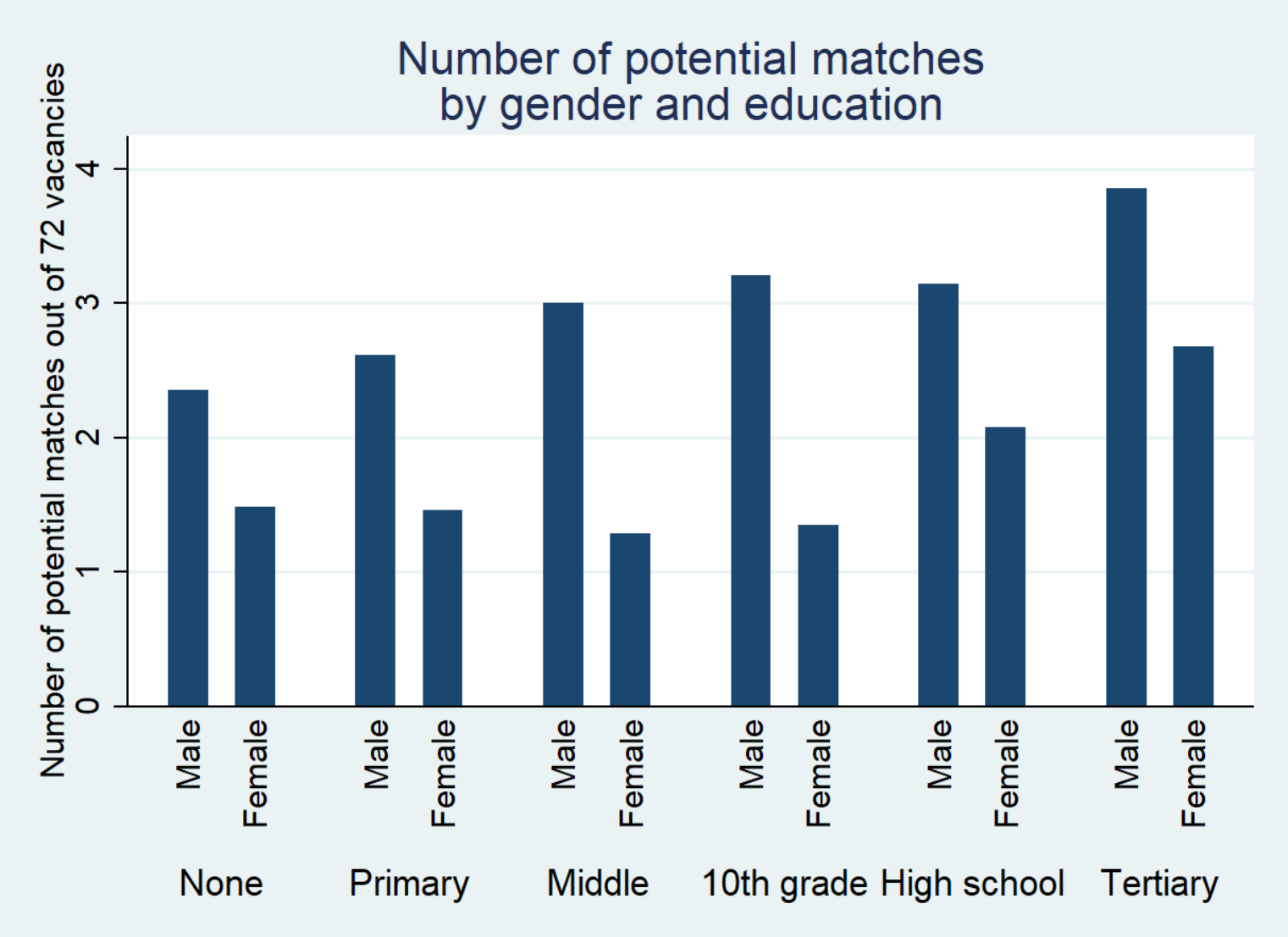
Most firms have no women...



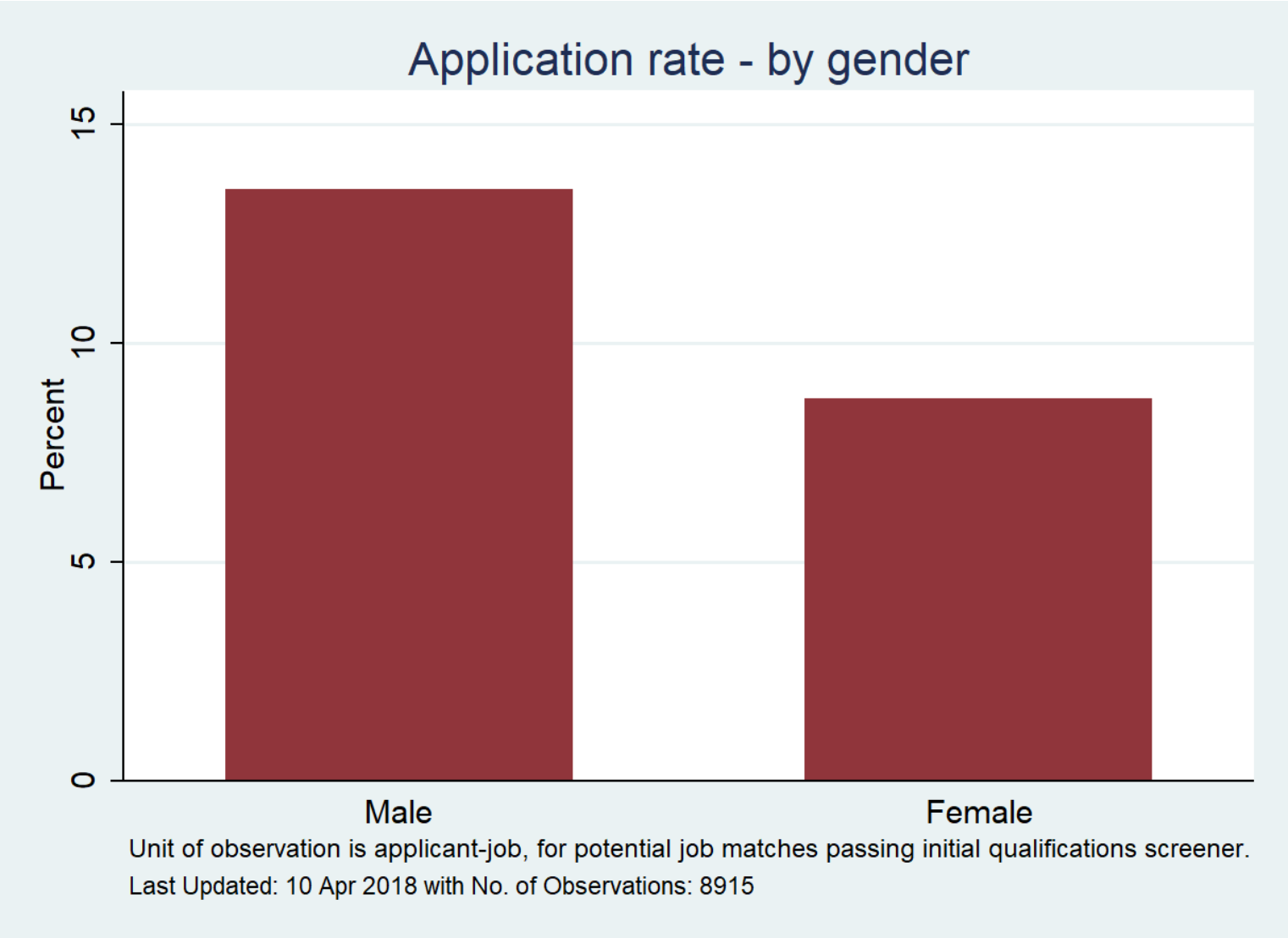
... but are willing to hire them



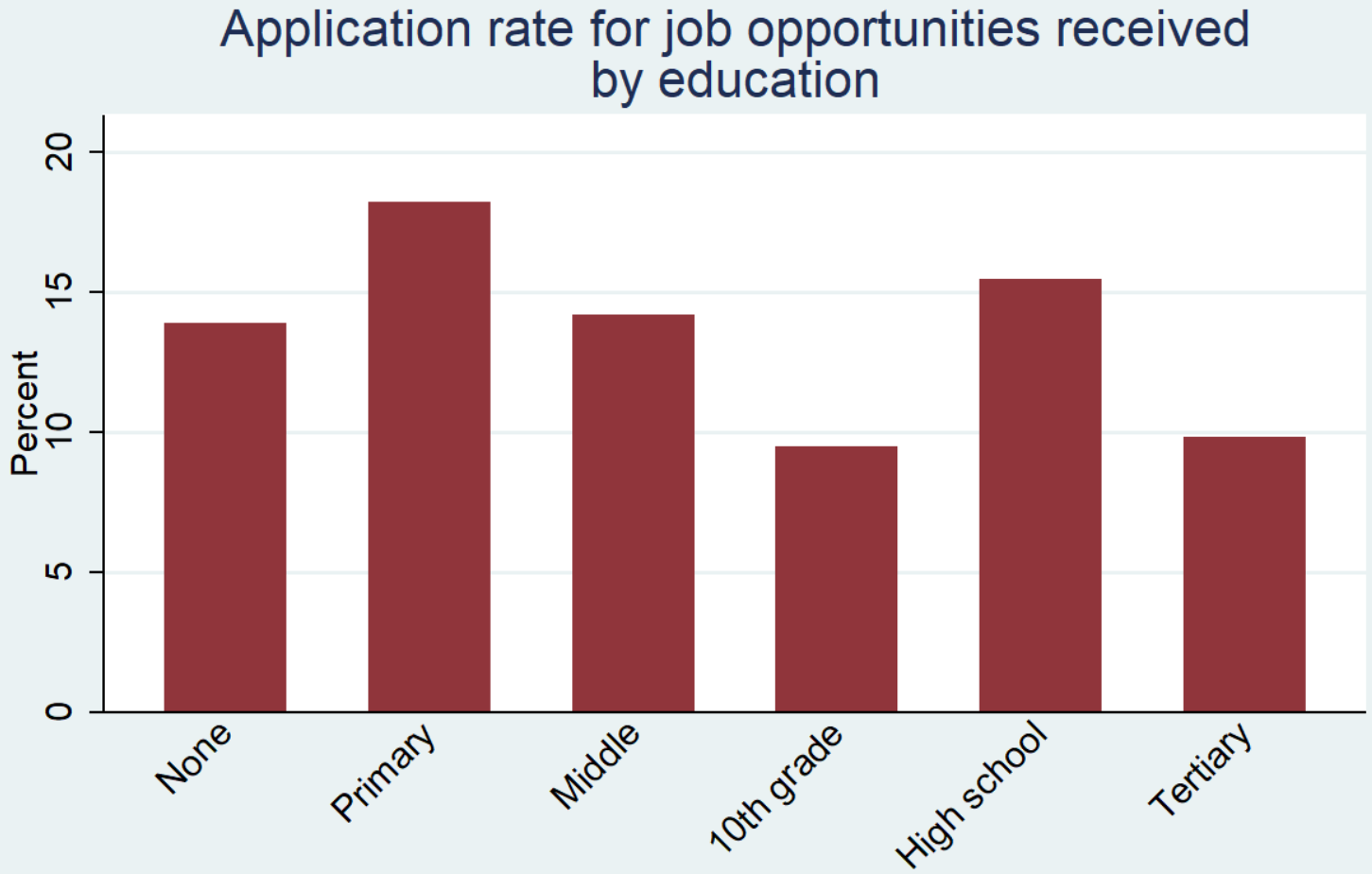
Potential matches across gender and education



Subscribers apply to 15% of proposed matches



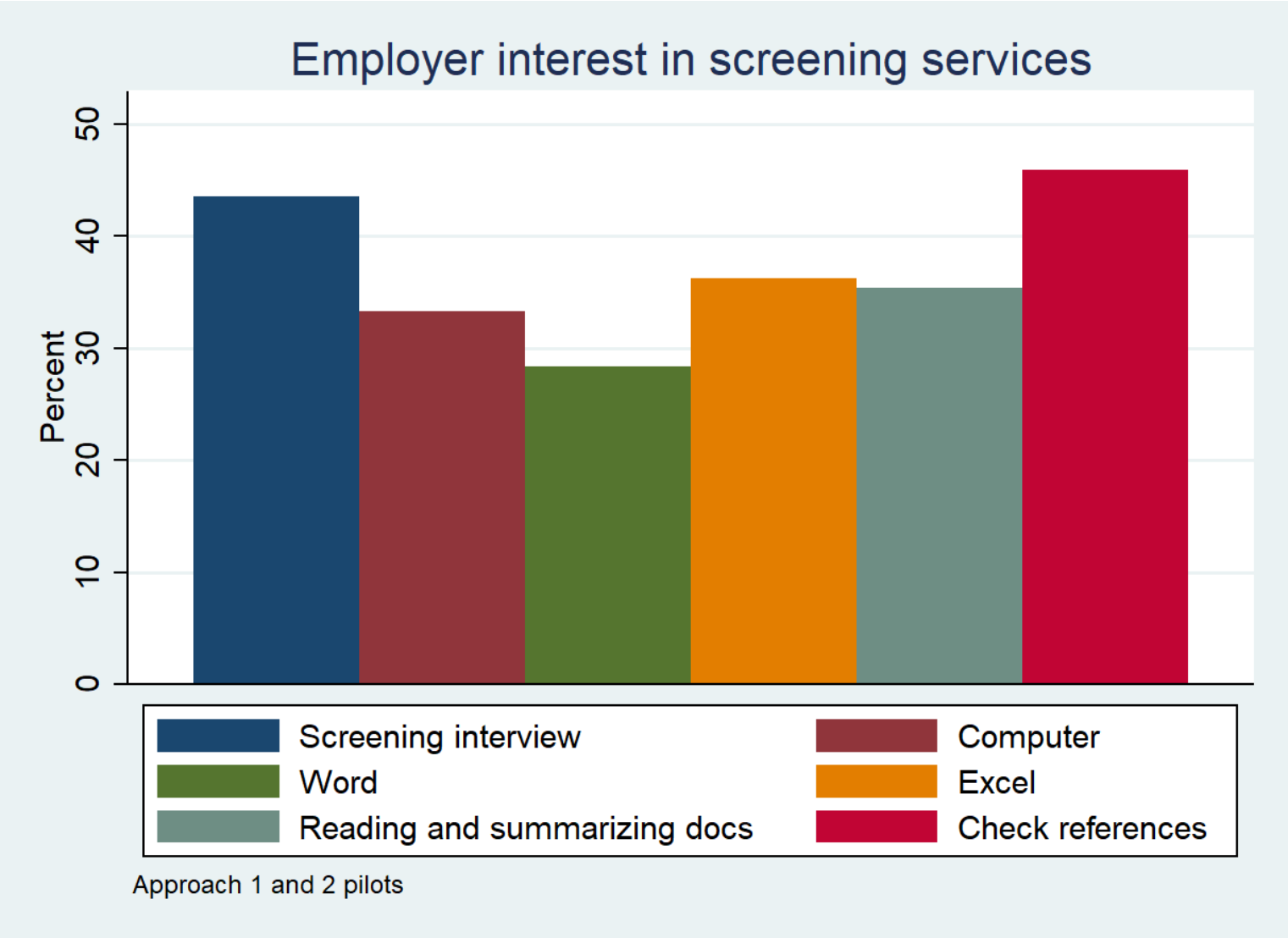
Subscribers apply to 15% of proposed matches



Screening and skills testing

- Does the intervention **improve matching** or do treated women displace other similar women?
- Firm randomization: test effects on **hiring and composition**
- Offer free HR services: **test and score** applicants
- Test directly for effects on **pool quality**
- **91% of employers** surveyed express interest in screening service

Screening and skills testing



Transport intervention offer through Job Talash

- **Transport RCT underway**: offers sent for first 50 jobs to pilot subsets of T1, T2 and C
- Initial results soon on **applications** and **hiring**

Transport intervention offer through Job Talash

Scale up - test impacts on:

- **Applications**
- **Hiring**
- **Location** - apply to and take up work further from home?
- **Composition** for firms
- **Pool quality** (with testing services)
- **Norms**: acceptability of women's work, mobility, gender segregation

Supplementary slides

Takeup of green line for daily commute

	(1)	(2)
Distance to closest built stop (T1)	-0.001***	-0.003***
	(0.000)	(0.001)
Distance to closest stop (T1 / T2 / C)	0.000	0.000
	(0.001)	(0.001)
Distance to closest built stop sq		0.000***
		(0.000)
Distance to closest stop sq		-0.000
		(0.000)
Constant	0.146***	0.047
	(0.051)	(0.054)
Observations	48106	48106
Sample mean dependent variable		0.010

Each observation is one adult HH member in year 2015-6. All specifications include controls for female, age, age squared, years of education and years of education squared.

Matching variables

- Distance to center
- Punjab Directory of Industries: Number of manufacturing firms, weighted by distance; total firm investment, weighted by distance
- Population density
- 2010 government survey: Income; Education; Trip cost and duration; Years at location; Owns house; Rent; House area; Number of rooms; Monthly transport expenditure; Vehicle ownership; HH size
- To address **pre-trends**, 1998 census: education (primary male, primary female, matric male, matric female), demographics (proportion age 10 or older, 18 or older), proportion religious minorities
- HH and individual variables for **neighboring** zones

Matching method

Construct Mahalanobis distance on vector of baseline characteristics between each C zone and corresponding potential T1 zones:

$$D_M(x) = \sqrt{(x_i - x_j)' S^{-1} (x_i - x_j)}$$

Where

- x_i and x_j are baseline characteristics of a C and T1 zone
- S is their covariance matrix
- Select pairs of C and T1 zones with a given maximum D_M
- Repeat for C-T2 matches
- Select C zones that have at least one matching T1 and one matching T2 zone
- Multiple matches allowed; use weights to correct for this

Back

Balance after matching (2010 government survey)

	T1 (line built)		T2 (line under construction)		
	Difference	SE	Difference	SE	Observations
Any income	0.00	(0.01)	-0.00	(0.01)	15140
Income	-282.09	(442.91)	340.28	(324.10)	15049
Ln income	-0.06	(0.08)	0.03	(0.07)	6347
Education: HS or less	0.05	(0.03)	0.01	(0.03)	15140
Education: HS	-0.01	(0.02)	0.01	(0.01)	15140
Education: higher	-0.04	(0.03)	-0.01	(0.02)	15140
Trip cost	-1.03	(3.18)	0.73	(2.64)	4754
Trip duration	115758.12	(204824.59)	64206.64	(161151.79)	4701
Years at location	2.51	(2.44)	4.75*	(2.43)	1012
Owns home	-0.03	(0.03)	-0.02	(0.03)	999
Rent	-259.83	(406.01)	-318.28	(383.56)	1012
House area	-1.93	(1.95)	-0.97	(1.94)	1005
Number of rooms	0.13	(0.32)	0.40	(0.30)	1012
Transport expenditure	-125.14	(371.24)	69.23	(306.55)	1003
HH income	-0.23	(0.48)	0.24	(0.32)	1012
Bicycle	-0.05	(0.09)	0.02	(0.07)	1012
Motorcycle	-0.04	(0.08)	0.08	(0.06)	1012
Number of members living in HH	0.08	(0.23)	0.15	(0.19)	1012
Number of members living away	0.03	(0.05)	0.06	(0.04)	1012

Back

Transit increases accessibility (time)

	(1)	(2)	(3)	(4)
Distance to closest built stop (T1)	3.6242***	6.1614***	-1.7914	-1.1902
	(0.7189)	(2.2718)	(3.5672)	(3.6828)
Distance to closest stop (T1 / T2 / C)	3.6000	-1.6908	6.4547	12.7046
	(3.2843)	(6.2372)	(9.5699)	(22.5379)
Distance to closest built stop sq		-0.1988	0.2220	0.1699
		(0.1591)	(0.2395)	(0.2495)
Distance to closest stop sq		1.4978	0.5908	-0.5643
		(0.8996)	(1.3424)	(2.6639)
Post			-32.1667***	-41.2925**
			(6.7696)	(19.8116)
Distance to closest built stop (T1) x post			7.9528***	8.1203***
			(1.8821)	(1.9628)
Distance to closest built stop sq x post			-0.4208***	-0.4354***
			(0.1141)	(0.1191)
Distance to closest stop x post			-8.1456	-0.2845
			(4.8810)	(12.2757)
Distance to closest stop sq x post			0.9070	-0.1952
			(0.7362)	(1.4676)
Observations	529	529	1567	1567
Donut FE	No	No	No	Yes

Dependent variable is the real estate agent's report of travel time by public transport to central

Transit increases accessibility (fare)

	(1)	(2)	(3)	(4)
Distance to closest built stop (T1)	3.1740***	5.1832***	-1.5428	-1.4295
	(0.3146)	(0.7526)	(1.6676)	(1.6709)
Distance to closest stop (T1 / T2 / C)	-1.5227	-3.9916*	-4.1407	-5.6932
	(1.4657)	(1.9926)	(3.9436)	(9.8842)
Distance to closest built stop sq		-0.1562***	0.1817	0.1723
		(0.0529)	(0.1260)	(0.1263)
Distance to closest stop sq		0.7419**	0.6631	0.8013
		(0.3020)	(0.5403)	(1.2703)
Post			-18.4958***	-62.6406***
			(4.8361)	(13.9527)
Distance to closest built stop (T1) x post			6.7259***	6.8417***
			(1.4400)	(1.4089)
Distance to closest built stop sq x post			-0.3379***	-0.3493***
			(0.1050)	(0.0977)
Distance to closest stop x post			0.1491	25.4262***
			(3.6806)	(9.3578)
Distance to closest stop sq x post			0.0788	-2.8363**
			(0.5113)	(1.2006)
Observations	526	526	1551	1551
Donut FE	No	No	No	Yes

Dependent variable is the real estate agent's report of total fare for public transport to central

Transit increased overall public transport use

Commutes by public transport (conditional on commute)					
	(1)	(2)	(3)	(4)	(5)
Time to access central Lahore through public transport	-0.0017***	-0.0013***	-0.0005**	-0.0006*	-0.0004**
	(0.0005)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Observations	7261	7181	13182	10776	9141
Additional control variables	No	Yes	Yes	Yes	Yes
Donut FE	No	Yes	Yes	Yes	Yes
Geographic sample	Full	Full	Full	T1 T2	T1 C
Specification	XS	XS	Panel	Panel	Panel
Hansen's J p-value	0.9405	0.9949	0.2105	0.1921	0.8069
First-stage F-stat	11.9319	19.9707	20.0149	18.3624	39.0177
Sample mean pre	0.0700	0.0700	0.0700	0.0700	0.0800

Results shown for second stage of 2SLS estimate. Each observation is one adult-round. Households and individuals who moved into their current residence after the Green Line was built are excluded. The first panel shows estimations on the full sample, with observations taking on a zero value for adults who do not commute. The second panel shows estimations on the sample of adults who report a commute only. The probability of commuting does not change (Table **??**), so interpretation of the results conditional on commuting is not affected by sample selection concerns. Standard errors clustered at the zone level. * $p < .1$, ** $p < .05$, *** $p < .01$.

Most switching from motorbikes

	Motorcycle				
Time to access central Lahore by public transport	0.0006*	0.0004*	0.0003***	0.0003**	0.0003***
	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Observations	30487	30103	59846	48854	41270
Controls	No	Yes	Yes	Yes	Yes
Donut FE	No	Yes	Yes	Yes	Yes
Geographic sample	Full	Full	Full	T1 T2	T1 C
Specification	XS	XS	Panel	Panel	Panel

Results shown for second stage of 2SLS estimate. Each observation is one adult-round. Households and individuals who moved into their current residence after the Green Line was built are excluded. All panels show estimations on the full sample, with observations taking on a zero value for adults who do not commute. Standard errors clustered at the zone level. * $p < .1$, ** $p < .05$, *** $p < .01$.

Approximate calculations of averted emissions from switching to public transport

Parameter	Estimate	Source	
Number of switchers	35000	HH survey regression estimates*	
Mean travel distance (daily round trip)	9	Baseline HH survey (2010 HIS)	
Proportion switching completely to public transport	0.55	Descriptive data from rider survey	
Proportion of trip on private modes for mixed mode trips	0.5	Assumption	
Gallons gasoline equivalent averted			
	per mile traveled on bus instead of motorcycle	0.01	(1/f) - (1 / c)
Passenger-miles traveled on bus instead of motorcycle	228375	(1 - i*j) * g * h	
Gallons gasoline equivalent averted - total	2,282	k*l	
Tons CO2 averted per year	5,914	o * 264 working days	

Increase in activity of small local businesses - **work in progress**

	(1)	(2)	(3)	(4)
	Number of businesses		Business type Herfindahl	
Public transport travel time to central Lahore	-0.013***		0.001	
	(0.003)		(0.000)	
Constructed public transport travel time		-0.012***		0.000
		(0.004)		(0.001)
Observations	310	598	310	599
Donut FE	Yes	Yes	Yes	Yes
Specification	XS	XS	XS	XS
Sample	T1 C	Full	T1 C	Full
Hansen's J p-value	0.287	0.278	0.908	0.518
First-stage F-stat	53.568	8.258	53.568	8.186
Sample mean (C)		1.6		

Robust standard errors clustered by zone shown in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Increase in density - **work in progress**

	(1) Density
Distance to T1 stop	-0.069***
	(0.002)
Post	0.402***
	(0.005)
Distance to T1 stop x post	-0.011***
	(0.000)
Distance to T2 stop	-0.003***
	(0.000)
Distance to T2 stop x post	-0.006***
	(0.000)
Distance to C stop	-0.058***
	(0.006)
Distance to C stop x post	0.006
	(0.003)
Observations	65618
"Donut" FE	Yes
Specification	Panel

Robust standard errors clustered by closest stop shown in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Real estate prices affected only where baseline density high - **work in progress**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln price		Ln rent		Ln price		Ln rent	
	Res	Com	Res	Com	Res	Com	Res	Com
Public transport travel	0.001	0.004	-0.002	0.005	-0.007***	-0.007*	-0.001	-0.006**
time to central Lahore	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)
Observations	334	334	346	346	599	547	610	590
Donut FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Specification	Panel - IV							
Sample	T1 C							
Hansen's J p	0.352	0.632	0.139	0.636	0.282	0.149	0.147	0.206
First-stage F	18.469	18.469	15.801	15.801	51.844	98.678	46.315	58.662
Baseline density	Low	Low	Low	Low	High	High	High	High

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

No effects on labor force participation - extensive margin

	(1)	(2)	(3)	(4)
	Work outside home			
Public transport travel time to central Lahore	-0.0000	0.0000	0.0000	0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Observations	35513	70419	57293	59458
Donut FE	Yes	Yes	Yes	Yes
Sample	Full	Full	T1 T2	Non-movers
Specification	XS	Panel	Panel	Panel
Hansen's J p-value	0.8963	0.8323	0.8183	0.6954
First-stage F-stat	18.2627	15.8145	19.7112	16.0686

Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$