

Online Appendix for “Measuring Productivity: Lessons from Tailored Surveys and Productivity Benchmarking”

David Atkin and Amit K. Khandelwal and Adam Osman*

January 11, 2019

A.1 Summary Statistics

We collected data for 219 rug producing firms in Fowa, Egypt between July 2011 and June 2014. We administered six rounds of surveys that captured information on rugs produced in the prior month including the rug specifications, prices and quantities of all inputs and outputs, labor hours spent on production and preparation activities. We also hired an independent quality assessor (a highly-skilled rug producer) who graded the rugs being produced at the time of the survey along eleven quality metrics (grading on a 1 to 5 scale, with 5 being the highest quality). Table A.1 provides the means and standard deviations of key variables used in our estimations.

After the last survey round, we set up a controlled lab in a rented space where all firms were asked to send their main rug producer to produce a rug with identical specifications using the same material inputs and capital equipment that we provided. The rug producer was paid a flat fee for his time. We recorded the rug’s final dimensions and the time taken to weave it. We also sent the rugs to be scored anonymously by both our quality assessor and a local professor of handicraft science. We use the average score for each quality metric in this paper. Table A.1 also reports the mean and standard deviations of the quality lab measures. Atkin, Khandelwal and Osman (2017) provides further details on the surveys and the sample.

Table A.2 reports the association between the rug specifications and the price of the rugs, overall output and total revenue of the firm during the month prior to the survey. The coefficients have signs consistent with our priors and the high R-squared suggests that specifications can explain much of the variation in these variables.

A.2 Survey-Based Productivity Measures

A.2.1 Quantity Production Functions

Our production function estimation follows Atkin, Khandelwal and Osman (2017). The first set of production function estimates do not control for rug specifications and hence provide our *unadjusted TFPQ* estimates. We estimate the following Cobb-Douglas production function:¹

$$x_{it} = \phi_{u,it} l_{it}^{\alpha_l} k_{it}^{\alpha_k} e^{\epsilon_{it}} \quad (1)$$

where x_{it} is the output (in m²) of firm i in period t , l_{it} is total labor hours, k_{it} is the number of active looms, and $\phi_{u,it}$ is the firm’s unadjusted TFPQ. The error term captures unanticipated shocks as well as omitted variables (the specifications of the rugs produced). To estimate the parameters of the production function, we take logs to obtain

$$\ln x_{it} = \ln \phi_{u,it} + \alpha_l^u \ln l_{it} + \alpha_k^u \ln k_{it} + \epsilon_{it} \quad (2)$$

The second set of production function estimates controls for rug specifications and provide our *specification-adjusted TFPQ* estimate. We estimate

$$\ln x_{it} = \ln \phi_{a,it} + \alpha_l^a \ln l_{it} + \alpha_k^a \ln k_{it} + \ln \lambda_{it}' \gamma + \epsilon_{it} \quad (3)$$

* Atkin: MIT, atkin@mit.edu. Khandelwal: Columbia Business School, ak2796@columbia.edu. Osman: University of Illinois, aosman@illinois.edu. We thank AbdelRahman Nagy and the Egypt field team.

¹We assume that output is Leontief in materials and therefore materials do not enter into the estimation.

where $\phi_{a,it}$ is the firm's specification-adjusted TFPQ and the vector λ_{it} includes six rug specifications: rug difficulty, thread count, thread type, number of colors, market segment, and narrow product type (where we include dummies for each value of the latter two categorical variables).² The error term now only captures unanticipated shocks and measurement error.

We estimate TFPQ via OLS and a control function. For the OLS regressions, we estimate (2) by regressing log of output on labor and capital. For (3), we add the six specifications to the regression. We estimate the production functions using the full set of duple firms in our sample of post-treatment rounds.³ Standard errors are clustered by firm. We report the estimates in columns 1 and 2 of Table A.3 below.

In the control function approach (Olley and Pakes (1996)) we assume capital is subject to adjustment costs, labor is a flexible input, and we use warp thread quantity as the proxy. We estimate the production functions using the one-step approach proposed by Wooldridge (2009), with l_{it-1} as the instrument for l_{it} , and cluster standard errors by firm. We report these estimates in columns 3 and 4 in Table A.3 below.

Unadjusted and specification-adjusted TFPQ are constructed from exponentiating the residuals of these production functions and then averaging across rounds for each firm.

A.2.2 Quality Production Functions

Quality productivity, TFPZ, is estimated as follows. As noted in the text, we begin by obtaining the consumers' valuations for quality implied by the following demand curve:

$$\ln x_{it} = (\sigma - 1) \sum_j \theta_j \ln q_{j,it} - \sigma \ln p_{it} + c_{it} \quad (4)$$

where q_j s are the eleven quality metrics, p is the price that firm i receives for its rug produced at the time of the survey, and c is a common price index. Using an estimate of $\sigma = 2.74$ from Broda and Weinstein (2006), we can re-write (4) as an estimating equation:

$$(\ln x_{it} + 2.74 \ln p_{it}) / (2.74 - 1) = \kappa + \sum_j \theta_j \ln q_{j,it} + \nu_{it} \quad (5)$$

where κ is a constant and ν is measurement error. The estimates of the θ_j s are reported in Table A.4.

With the estimates of θ in hand, we formulate the production function for producing consumers' valued quality, $\Pi_j q_j^{\hat{\theta}_j}$, with the same functional form as the quantity production function in (2):

$$\ln \left(\Pi_j q_{j,it}^{\hat{\theta}_j} \right) = \ln \zeta_{u,it} + \beta_l^u \ln l_{it} + \beta_k^u \ln k_{it} + \epsilon_{it} \quad (6)$$

As before, we can estimate (6) via OLS or a control function to obtain *unadjusted TFPZ*. The results are reported in Table A.5.

Analogously to specification-adjusted TFPQ, we can recover *specification-adjusted TFPZ* by controlling for specifications in the quantity production production:

$$\ln \left(\Pi_j q_{j,it}^{\hat{\theta}_j} \right) = \ln \zeta_{a,it} + \beta_l^a \ln l_{it} + \beta_k^a \ln k_{it} + \ln \lambda'_{it} \delta + \epsilon_{it} \quad (7)$$

The results of estimating (7) via OLS and a control function are reported in Table A.5.

Unadjusted and specification-adjusted TFPZ are constructed from exponentiating the residuals of these production functions and then averaging across rounds for each firm.

A.2.3 Capabilities Production Functions

For unadjusted firm capabilities, which we term *unadjusted TFPC*, we multiply output by the quality aggregator to formulate a combined production function for $x_{it} \Pi_j q_{j,it}^{\hat{\theta}_j}$, the combination of quantity and quality that consumers value in their utility function.

²As discussed in Atkin, Khandelwal and Osman (2017), we have two samples of firms that we pool over in this production function estimation. For the firms in the first sample, we did not record the market segment or rug difficulty. We replace these missing values with the corresponding values from the subsequent survey round.

³This differs from Atkin, Khandelwal and Osman (2017) where we estimate the production function only on the sample of control firms to avoid having to take a stance on the Markov process governing productivity evolution over time for the treatment firms. In this paper, since we are simply interested in cross-sectional comparisons, we only focus on the post-treatment sample where export status is not changing and estimate the production function over all firms.

$$\ln \left(x_{it} \Pi_j q_{j,it}^{\hat{\theta}_j} \right) = \ln \zeta_{u,it} + \ln \phi_{u,it} + (\alpha_l^u + \beta_l^u) \ln l_{it} + (\alpha_k^u + \beta_k^u) \ln k_{it} + \epsilon_{it} \quad (8)$$

As before, we estimate (8) via OLS and a control function, and report the results in Table A.6. The structure of the production function implies that the coefficients of the capabilities production function equal the sum of the coefficients from the quantity and quality production functions (e.g., the sum of the labor coefficient in column 1 of Table A.3 and the labor coefficient in column 1 of Table A.5.⁴). Unadjusted TFPQ is the product $\zeta_u \phi_u$.

Similarly, we can estimate *specification-adjusted TFPQ* from the following production function:

$$\ln \left(x_{it} \Pi_j q_{j,it}^{\theta_j} \right) = \ln \zeta_{a,it} + \ln \phi_{a,it} + (\alpha_l^a + \beta_l^a) \ln l_{it} + (\alpha_k^a + \beta_k^a) \ln k_{it} + \ln \lambda'_{it} (\gamma + \delta) + \epsilon_{it} \quad (9)$$

with the results reported in A.6. Specification-adjusted TFPQ is the product $\zeta_a \phi_a$.

A.2.4 Revenue Production Functions

We estimate a revenue production function using the following specification:

$$\ln r_{it} = \ln TFPR_{it} + \eta_l \ln w_{it} + \eta_k \ln rk_{it} + \epsilon_{it} \quad (10)$$

where r_{it} is the revenue of the firm, w_{it} is the wage bill, and rk_{it} is the value of the capital stock. We estimate (10) via OLS and a control function and report the results in A.7. (Note that we do not control for specifications in these regressions). TFPR is constructed from exponentiating the residual of this production function and then averaging across rounds for each firm.

A.3 Description of Appendix Figures and Tables

- Table A.1 provides summary statistics for the variables used to estimate the production functions.
- Table A.2 estimates the relationship between the rug specifications and price, output and revenue.
- Table A.3 reports the coefficients from the quantity production function.
- Table A.4 reports the θ s coefficients from the demand estimation.
- Table A.5 reports the coefficients from the quality production function.
- Table A.6 reports the coefficients from the capabilities production function.
- Table A.8 is the correlation matrix for the measures used in the paper estimated using OLS.
- Table A.9 is the correlation matrix for the measures used but estimated using a control function approach.
- Table A.10 shows the correlation matrix including both OLS and control function values.
- Figure A.1 reports the distribution of the three TFPQ measures: unadjusted TFPQ (ϕ_u), specification-adjusted TFPQ (ϕ_a) and Lab TFPQ. We plot the OLS version of each productivity measure (scaled by dividing through by the average).
- Figure A.2 reports the distribution of the three TFPZ measures: unadjusted TFPZ (ζ_u), specification-adjusted TFPZ (ζ_a) and Lab TFPZ. We plot the OLS version of each productivity measure (scaled by dividing through by the average).
- Figure A.3 reports the distribution of the three TFPC measures: unadjusted TFPC ($\zeta_u \phi_u$), specification-adjusted TFPC ($\zeta_a \phi_a$) and Lab TFPC. We plot the OLS version of each productivity measure (scaled by dividing through by the average).
- Figure A.4 reports the distribution of TFPR. We plot the OLS version (scaled by dividing through by the average).

⁴Due to missing observations, the coefficients do not line up exactly. The paper uses the TFPC estimate that comes from the product of the individual ζ_u and ϕ_u estimates.

References

- Atkin, David, Amit K. Khandelwal, and Adam Osman.** 2017. “Exporting and Firm Performance: Evidence from a Randomized Experiment.” *The Quarterly Journal of Economics*, 132(2): 551–615.
- Broda, Christian, and David E. Weinstein.** 2006. “Globalization and the Gains From Variety.” *The Quarterly Journal of Economics*, 121(2): 541–585.
- Olley, G Steven, and Ariel Pakes.** 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6): 1263–97.
- Wooldridge, Jeffrey M.** 2009. “On estimating firm-level production functions using proxy variables to control for unobservables.” *Economics Letters*, 104(3): 112–114.

A.4 Appendix Tables and Figures

Table A.1: Summary Statistics

	Mean	Standard Deviation	Observations
Output (Square Meters)	59.43	(75.04)	900
Labor Hours	5.55	(0.29)	900
Capital (Looms)	0.08	(0.27)	912
(log) Thread Quantity	7.46	(0.28)	913
Difficulty Control	3.23	(0.83)	926
(log) Number of Colors	1.72	(0.99)	914
Mid-Market Segment=1	0.23	(0.42)	923
Low-Market Segment=1	0.42	(0.49)	923
Price (EGP/Square Meter)	29.29	(47.27)	913
Survey Quality: Packedness	3.25	(0.86)	913
Survey Quality: Corners	3.14	(0.85)	913
Survey Quality: Waviness	3.15	(0.84)	913
Survey Quality: Weight	3.22	(0.84)	913
Survey Quality: Touch	3.19	(0.49)	913
Survey Quality: Warp Thread Tightness	3.18	(0.81)	913
Survey Quality: Firmness	3.02	(0.56)	913
Survey Quality: Design Accuracy	3.32	(0.86)	913
Survey Quality: Ward Thread Packedness	3.19	(0.83)	913
Survey Quality: Inputs	3.20	(0.87)	913
Survey Quality: Loom	2.04	(0.24)	913
Lab Quality: Packedness	3.34	(0.63)	187
Lab Quality: Corners	3.29	(0.63)	187
Lab Quality: Waviness	3.28	(0.60)	187
Lab Quality: Weight	3.60	(0.83)	187
Lab Quality: Touch	3.29	(0.50)	187
Lab Quality: Warp Thread Tightness	2.95	(0.66)	187
Lab Quality: Firmness	3.24	(0.65)	187
Lab Quality: Design Accuracy	3.46	(0.62)	187
Lab Quality: Ward Thread Packedness	3.27	(0.68)	187
Lab Quality: Inputs	4.00	(0.00)	187
Lab Quality: Loom	2.00	(0.00)	187

Notes: Table reports summary statistics of the variables used to estimate the production functions. “Quality” denotes the 11 quality metrics. “Lab” denotes the quality metrics from the controlled lab, which are averaged over grades given by the quality assessor and professor of handicraft science. “EGP” denotes Egyptian pounds (which was around 6.31 pounds to one USD over the sample period). See Atkin et al (2017) for more details about the sample and variables.

Table A.2: Outcomes and Specifications

	(1)	(2)	(3)
	Price	Output	Revenue
(log) Thread Quantity	0.11 (0.14)	-0.01 (0.12)	0.11 (0.11)
Difficulty Control	0.13*** (0.03)	-0.06* (0.03)	0.07** (0.03)
(log) Number of Colors	-0.02 (0.03)	-0.05* (0.03)	-0.06** (0.03)
Low-Market Segment=1	-0.84*** (0.08)	0.52*** (0.07)	-0.30*** (0.07)
Mid-Market Segment=1	-0.60*** (0.08)	0.31*** (0.08)	-0.26*** (0.06)
Product Type Dummies (6 Categories)	Yes	Yes	Yes
Thread Type Dummies (6 Categories)	Yes	Yes	Yes
r2	.536	.454	.117
N	825	890	818

Notes: Table reports the results of estimating the log price, log output and log revenue on the six specifications. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.3: Quantity Production Function

	(1)	(2)	(3)	(4)
	Unadjusted (OLS)	Adjusted (OLS)	Unadjusted (CF)	Adjusted (CF)
Labor	0.61*** (0.11)	0.65*** (0.09)	1.41** (0.70)	1.31*** (0.51)
Capital Inputs	0.19* (0.11)	0.24** (0.10)	0.41** (0.19)	0.24** (0.12)
(log) Thread Quantity		-0.02 (0.11)		-0.28* (0.17)
Difficulty Control		-0.06** (0.03)		-0.12*** (0.04)
(log) Number of Colors		-0.05* (0.03)		-0.07** (0.03)
Low-Market Segment=1		0.56*** (0.07)		0.55*** (0.08)
Mid-Market Segment=1		0.37*** (0.07)		0.34*** (0.08)
Product Type Dummies (6 Categories)	No	Yes	No	Yes
Thread Type Dummies (6 Categories)	No	Yes	No	Yes
r2	.046	.506	.000	.508
N	899	889	595	591

Notes: Table reports the results of estimating the quantity production function. Columns 1 and 3 estimate the unadjusted production function. Columns 2 and 4 estimate the specification-adjusted production function. Columns 1-2 estimate via OLS and columns 3-4 estimate via a control function. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.4: Consumers' Valuation of Quality (θ 's)

	(1) Consumer Quality Valuation
Packedness	0.18 (0.26)
Corners	-0.10 (0.26)
Waviness	-0.09 (0.26)
Weight	-0.10 (0.22)
Touch	0.15 (0.27)
Warp Thread Tightness	0.87*** (0.25)
Firmness	-0.31 (0.33)
Design Accuracy	0.76*** (0.20)
Ward Thread Packedness	0.51** (0.24)
Inputs	-0.09 (0.23)
Loom	-0.70* (0.41)
r2	.168
N	892

Notes: Table reports the results of estimating the demand curve to obtain consumers' valuation of quality, θ_j s. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.5: Quality Production Function

	(1)	(2)	(3)	(4)
	Unadjusted (OLS)	Adjusted (OLS)	Unadjusted (CF)	Adjusted (CF)
Labor	0.07 (0.04)	-0.01 (0.03)	-0.18 (0.34)	-0.05 (0.12)
Capital Inputs	0.01 (0.05)	0.08** (0.03)	-0.02 (0.08)	0.09* (0.05)
(log) Thread Quantity		0.02 (0.03)		0.08 (0.06)
Difficulty Control		0.37*** (0.01)		0.35*** (0.02)
(log) Number of Colors		0.02** (0.01)		0.02 (0.01)
Low-Market Segment=1		-0.07*** (0.02)		-0.10*** (0.03)
Mid-Market Segment=1		-0.07*** (0.03)		-0.09*** (0.03)
Product Type Dummies (6 Categories)	No	Yes	No	Yes
Thread Type Dummies (6 Categories)	No	Yes	No	Yes
r ²	.002	.672	.052	.742
N	891	882	589	585

Notes: Table reports the results of estimating the quality production function. Columns 1 and 3 estimate the unadjusted production function. Columns 2 and 4 estimate the specification-adjusted production function. Columns 1-2 estimate via OLS and columns 3-4 estimate via a control function. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.6: Capabilities Production Function

	(1)	(2)	(3)	(4)
	Unadjusted (OLS)	Adjusted (OLS)	Unadjusted (CF)	Adjusted (CF)
Labor	0.67*** (0.10)	0.63*** (0.09)	1.32** (0.65)	1.12*** (0.35)
Capital Inputs	0.20* (0.10)	0.32*** (0.10)	0.35** (0.15)	0.28** (0.14)
(log) Thread Quantity		0.00 (0.11)		-0.19 (0.19)
Difficulty Control		0.30*** (0.03)		0.23*** (0.04)
(log) Number of Colors		-0.02 (0.03)		-0.05 (0.03)
Low-Market Segment=1		0.47*** (0.07)		0.43*** (0.08)
Mid-Market Segment=1		0.29*** (0.08)		0.24*** (0.08)
Product Type Dummies (6 Categories)	No	Yes	No	Yes
Thread Type Dummies (6 Categories)	No	Yes	No	Yes
r2	.062	.341	.005	.279
N	891	882	589	585

Notes: Table reports the results of estimating the capability production function. Columns 1 and 3 estimate the unadjusted production function. Columns 2 and 4 estimate the specification-adjusted production function. Columns 1-2 estimate via OLS and columns 3-4 estimate via a control function. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.7: Revenue Production Function

	(1)
	Log Revenue
Wage Bill	0.46*** (0.10)
Value of Captial Stock	0.08** (0.04)
r2	.070
N	794

Notes: Table reports the results of estimating the revenue production function. Column 1 estimates via OLS and column 2 estimates via a control function. Standard errors clustered at the firm level in parentheses. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.8: Correlation Matrix (OLS)

	Lab TFPQ	Lab TFPC	Lab TFPZ	Unadj. TFPQ	Adj. TFPQ	TFPR	Unadj. TFPZ	Adj. TFPZ	Unadj. TFPC
Lab TFPC	0.40***	1.00							
Lab TFPZ	0.07	0.94***	1.00						
Unadj. TFPQ	0.02	-0.15**	-0.15**	1.00					
Adj. TFPQ	0.14*	0.15**	0.11	0.42***	1.00				
TFPR	-0.02	0.03	0.03	0.10	0.28***	1.00			
Unadj. TFPZ	-0.05	0.45***	0.50***	-0.40***	0.01	0.07	1.00		
Adj. TFPZ	-0.17**	0.14*	0.22***	0.08	0.15**	-0.05	0.52***	1.00	
Unadj. TFPC	0.03	0.10	0.11	0.78***	0.71***	0.20***	0.07	0.40***	1.00
Adj. TFPC	0.07	0.19***	0.19**	0.38***	0.94***	0.24***	0.18***	0.44***	0.77***

Notes: Table reports the correlation between the variable at the top of each column with the variable in the associated row. Significance: * 0.10, ** 0.05, *** 0.01.

Table A.9: Correlation Matrix (Control Function)

	Lab TFPQ	Lab TFPC	Lab TFPZ	Unadj. TFPQ	Adj. TFPQ	TFPR (CF)	Unadj. TFPZ	Adj. TFPZ	Unadj. TFPC
Lab TFPC	0.40***	1.00							
Lab TFPZ	0.07	0.94***	1.00						
Unadj. TFPQ (CF)	0.03	-0.15**	-0.16**	1.00					
Adj. TFPQ (CF)	0.16**	0.18**	0.14*	0.45***	1.00				
TFPR (CF)	0.01	0.04	0.02	0.07	0.22***	1.00			
Unadj. TFPZ (CF)	-0.04	0.45***	0.50***	-0.40***	0.03	0.11	1.00		
Adj. TFPZ (CF)	-0.17**	0.13*	0.21***	0.18***	0.16**	-0.08	0.48***	1.00	
Unadj. TFPC (CF)	0.04	0.09	0.09	0.81***	0.70***	0.17**	0.02	0.44***	1.00
Adj. TFPC (CF)	0.10	0.22***	0.20***	0.45***	0.94***	0.19***	0.17**	0.43***	0.77***

Notes: Table reports the correlation between the variable at the top of each column with the variable in the associated row. Significance: * 0.10, ** 0.05, *** 0.01.

Figure A.1: Distribution of TFPQ

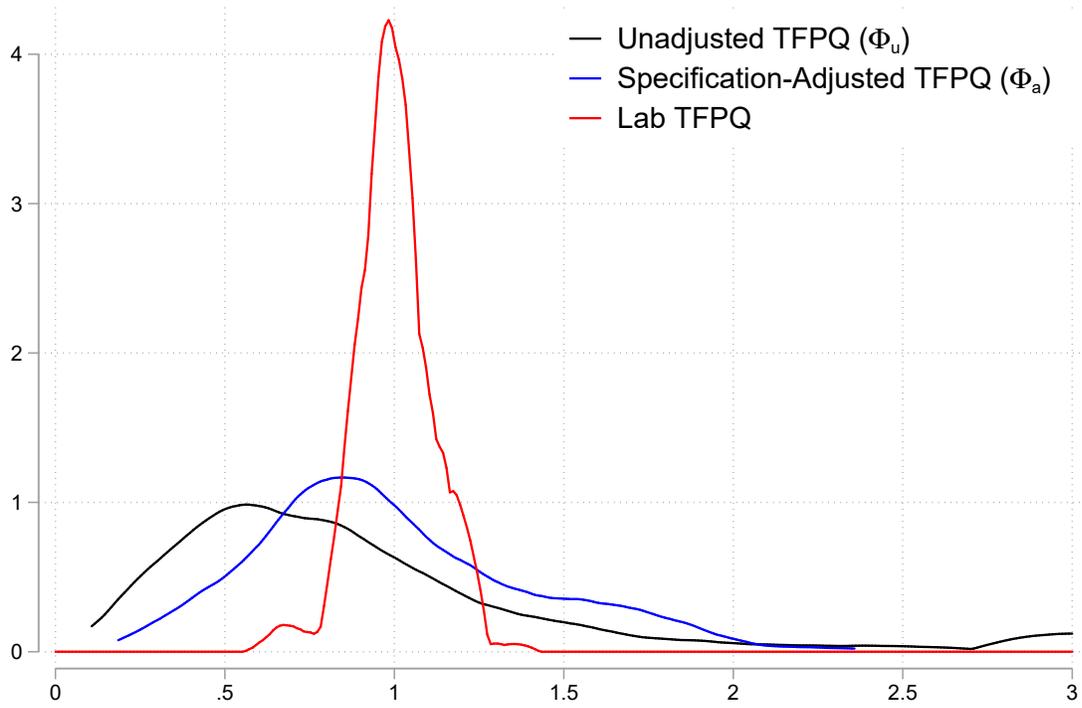


Figure A.2: Distribution of TFPZ

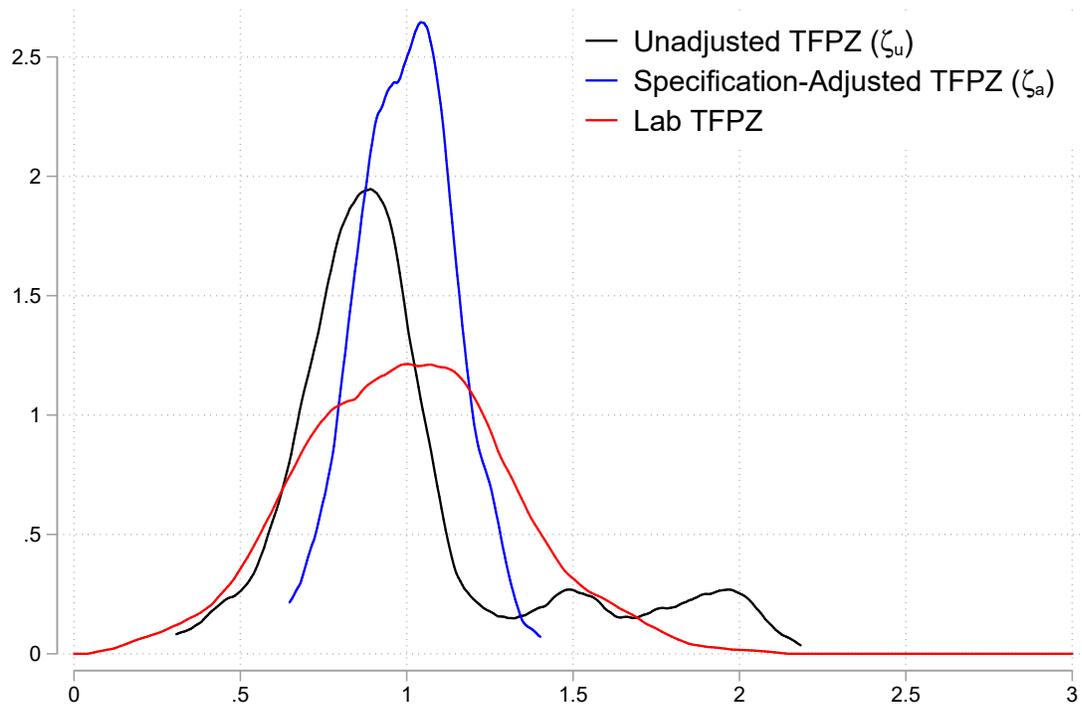


Figure A.3: Distribution of TFPC

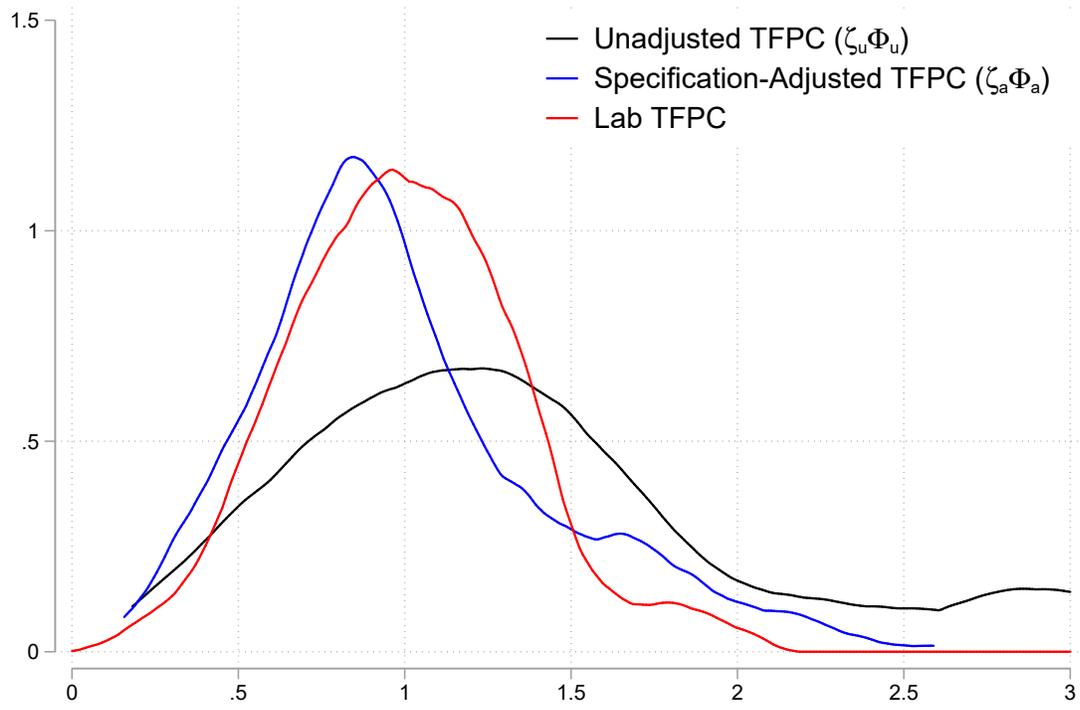


Figure A.4: Distribution of TFPR

