Particulate Pollution and the Productivity of Pear Packers:

Online Appendix to the Corrigendum

Tom Chang

Joshua S. Graff Zivin

Tal Gross

Matthew J. Neidell

As detailed in the published corrigendum, our 2016 paper involved an error. We thought that the pollution data in our possession was based on 6-day averages of $PM_{2.5}$, but the data actually consisted of 1-in-6-day measures covering 24 hours. In the published corrigendum, we present a replication of our main results – Table 3, Figure 4, and Figure 5 – using only "measurement days," days on which we observe $PM_{2.5}$.

This online appendix presents two sets of additional results. First, we present all other tables and figures from the original paper that are in any way altered by the restriction to measurement days. Second, we present all tables and figures using a more parsimonious specification. As detailed in the corrigendum, the decrease in sample size from only using measurement days hinders our ability to include the full set of fixed effects from the original specification, and leads to spurious results for the non-linear model specifications. In the more-parsimonious model specification, we replace year-by-month fixed effects with separate year and month fixed effects, and we replace day-of-week fixed effects with a weekend fixed effect. Our rationale for these changes is straightforward. While PM and productivity both show seasonal and annual patterns, we only have 12 year-months in the sample. Therefore, we instead include separate fixed effects for year and month. Day-of-week fixed effects were allowed for the possibility that PM and productivity co-vary by each day of the week. While it is certainly true for the latter, any variation in pollution by day of week is arguably exogenous, with the possible exception of changes over the weekend. As such, we replace the six day-of-week dummy variables with a simple weekend dummy to account for the day of week pattern. This more-parsimonious specification removes the spurious patterns found in the full specification.

In the end, the linear results continue to suggest that $PM_{2.5}$ outdoors leads to a statistically and economically significant decrease in the productivity of indoor workers in our setting. Our conclusions regarding the non-linear results are more tempered. While we are reassured by the results from the more-parsimonious model specification, data limitations hamper our ability to fully explore this relationship.

This appendix proceeds as follows. The subsequent section reproduces the original paper's estimates when restricted to the measurement-day-only sample, with all tables labeled with the suffix "C" for "corrected." The model specification underlying each is identical to that in the original manuscript with two exceptions. First, the third and fourth columns of Table 2 in the original paper present logit models, but these logits fail to converge once we restrict the sample to measurement days, so those columns of Table 2C present linear probability models instead. Second, the censored quantile estimator in column 9 of Table 4C no longer converges when the sample is restricted to measurement days.

The next section presents results based on the more-parsimonious specification, with all tables labeled with the suffix "P" for "parsimonious." All specifications are identical to those in the original manuscript, except for the changes detailed above: the weekend fixed effect and the year and month fixed effects. One additional exception is the last column of Table 4P: the original logit model failed to converge and so we present a linear probability model instead. In addition, we have added a clarification to the table notes for Tables 2 and 4. The standard errors for all OLS models beyond the linear probability models are clustered on worker and date, while the standard errors for logit models, linear probability models, and quantile-regression models are clustered on date.

1. Estimates Based on the Measurement-Day-Only Sample

Table 1C. Sample Statistics

	Observations	Mean	Standard deviation	Minimum	Maximum
Panel A. Productivity variables					
Worked that day	1,330	0.94	0.24	0.00	1.00
Regular time hours per day	1,162	6.84	1.74	0.25	8.50
Regular time earnings per hour	1,162	6.99	2.83	0.08	17.17
Worked overtime that day	1,162	0.29	0.45	0.00	1.00
Overtime hours if overtime that day	332	1.72	1.04	0.25	8.50
Overtime hours per day	1,162	0.49	0.96	0.00	8.50
Overtime earnings per hour	332	12.25	5.75	0.14	29.70
Penalty	984	0.06	0.23	0.00	1.00
Panel B. Environmental variables					
$PM_{2.5}$ ($\mu g/m^3$)	33	9.49	7.00	1.90	39.70
$PM_{2.5}$ <10	21				
$PM_{2.5}$ 10–15	9				
$PM_{2.5}$ 15–20	1				
$PM_{2.5}$ 20–25	1				
$PM_{2.5}$ >25	1				
Ozone (ppb)	33	32.94	10.44	11.38	50.63
Nitrogen dioxide (ppb)	33	8.41	3.83	3.88	18.13
Carbon monoxide (ppm)	33	0.54	0.21	0.29	0.99
PM_{10}–$PM_{2.5}$	33	9.66	3.33	4.10	18.70
Dewpoint (degrees Fahrenheit)	33	9.93	3.65	-0.25	16.50
Rain (in)	33	0.03	0.17	0.00	1.00
Wind speed (mph)	33	4.18	1.24	2.31	8.69
Wind direction (from south)	33	0.52	0.51	0.00	1.00
Solar radiation/1,000 (Wh/m^2)	33	0.67	0.14	0.40	0.85
Temperature (degrees Fahrenheit)	33	75.00	8.93	54.95	90.05

Productivity variables consist of worker-day pear packer payroll records. Environmental variables consist of daily observations.

Table 2C. The Relationship between $PM_{2.5}$ and Labor Supply

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Working that day				Hours			
$PM_{2.5}$	- 0.005	0.026			0.023	0.025		
	(0.002)	(0.006)			(0.025)	(0.070)		
	[0.005]	[0.000]			[0.354]	[0.718]		
$PM_{2.5}$ 10–15			0.043	0.043			- 0.753	- 0.753
			(0.013)	(0.013)			(0.450)	(0.450)
			[0.002]	[0.002]			[0.095]	[0.095]
$PM_{2.5}$ 15–20			0.119	0.119			4.533	4.533
			(0.028)	(0.028)			(0.967)	(0.969)
			[0.000]	[0.000]			[0.000]	[0.000]
$PM_{2.5}$ 20–25			0.040	0.040			- 1.293	- 1.293
			(0.024)	(0.024)			(0.656)	(0.656)
			[0.108]	[0.109]			[0.049]	[0.049]
$PM_{2.5}$ >25			- 0.463				- 2.048	
			(0.025)				(0.750)	
			[0.000]				[0.006]	
R^2			0.128	0.047	0.707	0.701	0.731	0.725
N	1,166	1,129	1,330	1,293	1,162	1,143	1,162	1,143

Standard errors based on estimates clustered by date in parentheses for columns 1 through 4, standard errors based on estimates clustered by date and worker in parentheses for columns 5 through 8; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we observe $PM_{2.5}$. Columns 1 and 2 present marginal effects based on a logit model, columns 3 through 8 present results from ordinary least squares regressions. All regressions include wind speed, a wind direction dummy variable, dew point, a rain dummy variable, day-of-week dummy variables, and year-month dummy variables.

Table 4C. Robustness Checks

	Baseline estimates (1)	Exclude meterological controls (2)	Control flexibly for temperature (3)	Control for additional pollutants (4)	Add worker fixed effects (5)
$PM_{2.5}$	- 0.086	- 0.024	- 0.091	- 0.150	- 0.102
	(0.033)	(0.038)	(0.013)	(0.035)	(0.040)
	[0.008]	[0.528]	[0.000]	[0.000]	[0.011]
R^2	0.332	0.298	0.333	0.337	0.658
N	1,162	1,162	1,162	1,162	1,162

	Aggregate to six-day PM-measurement periods (6)	Median regression (7)	Minimum wage binds (8)	Censored median regression (9)	Low-quality packing (10)
$PM_{2.5}$	(Not applicable.)	- 0.082	- 0.008		0.035
		(0.022)	(0.004)		(0.002)
		[0.000]	[0.027]		[0.000]
R^2					
N		1,162	1,120	1,075	365

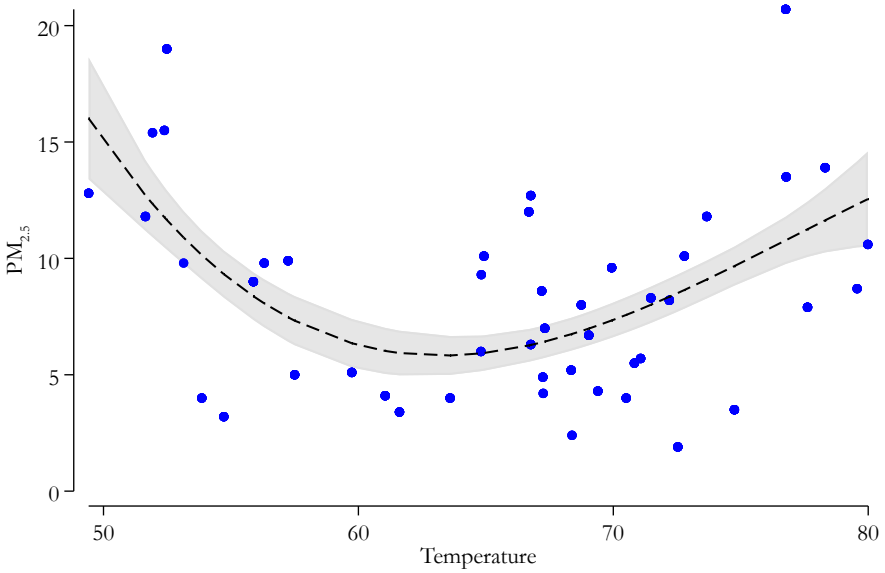
Standard errors based on estimates clustered by date and worker in parentheses for columns 1 through 5, standard errors based on estimates clustered by date in columns 6 through 10; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we observe $PM_{2.5}$. All regressions include day of week dummy variables and year-month dummy variables. All regressions except column 2 include wind speed, a wind direction dummy variable, dew point, and a rain dummy variable. Column 3 controls for temperature flexibly by including a series of indicator variables for each 5°F. Column 4 includes nitrogen dioxide, carbon monoxide, and coarse PM. In all regressions except for columns 8 and 10, the dependent variable is productivity during the regular-time shift, which is measured in earnings per hour. Column 8 uses whether the minimum wage binds as the dependent variable and column 10 uses "low-quality packing" as the dependent variable; both present marginal effects from a logit model.

Table 5C. The Relationship between $PM_{2.5}$ and Overtime Productivity

	Overtime hours (1)	Regular-time produtivity (2)	Overtime productivity (3)	Overtime productivity (4)
$PM_{2.5}$	- 0.030	- 0.144	- 0.071	0.015
	(0.021)	(0.024)	(0.021)	(0.040)
	[0.158]	[0.000]	[0.001]	[0.702]
R^2	0.484	0.295	0.295	0.413
N	1,162	332	332	332

Standard errors based on estimates clustered by date and worker in parentheses; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we obesrve $PM_{2.5}$. All regressions include wind speed, a wind direction dummy variable, dew point, a rain dummy variable, day-of-week dummy variables and year-month dummy variables. In columns 1 and 2, the dependent variable is the number of overtime hours worked. The dependent variable in column 3 is regular-time productivity and in columns 4 and 5 is overtime earnings, both limited to the sample of worker-days for which overtime hours exists. Productivity is measured in earnings per hour, though overtime productivity is deflated by 1.5 to account for time-and-a-half overtime pay.

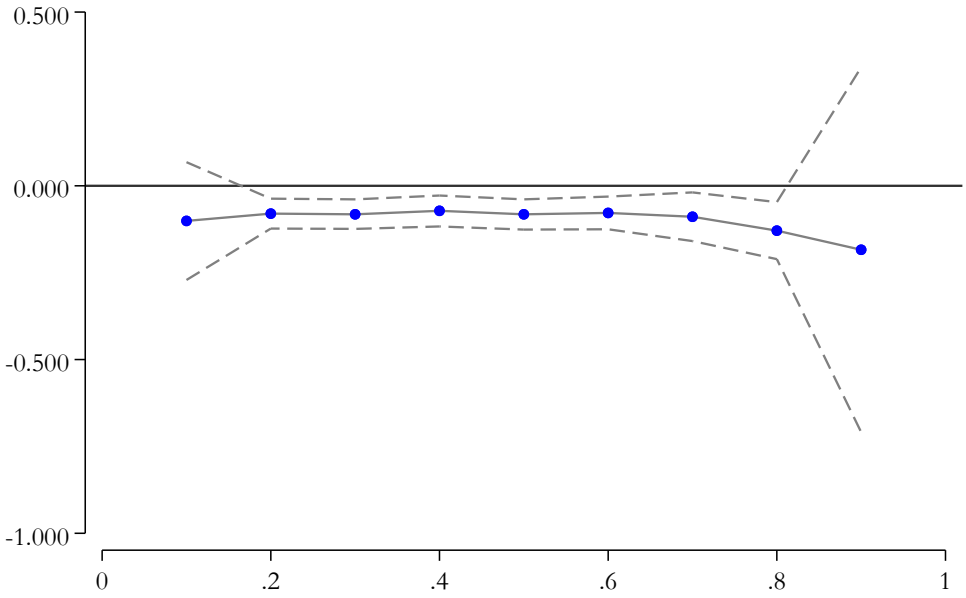
Figure 3C



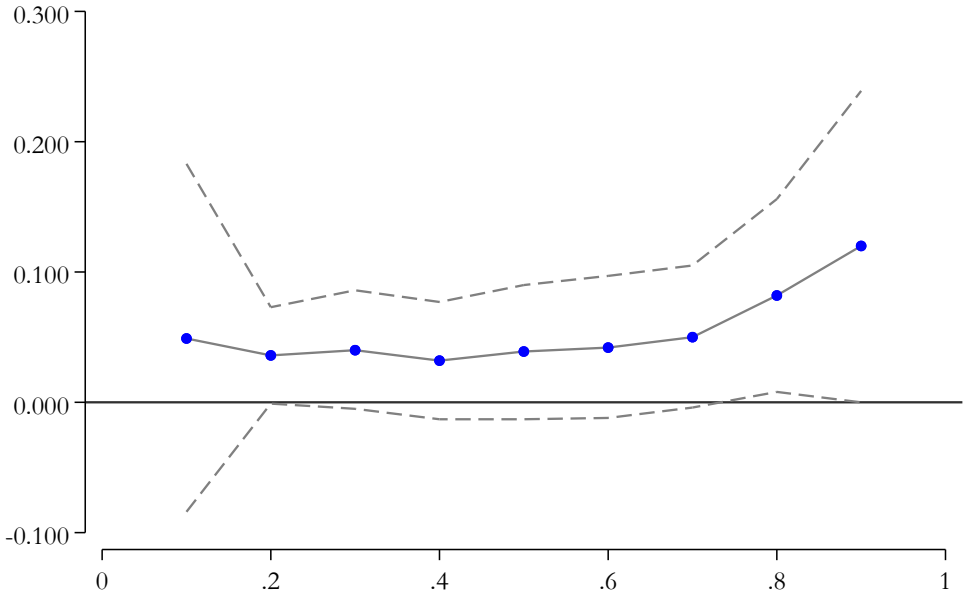
Notes: This figure presents $PM_{2.5}$ levels versus the average temperature in the date corresponding to that pollution observation. The solid line is the prediction based on a cubic series regression of $PM_{2.5}$ on temperature, with the shaded area indicating the 95-percent confidence intervals. The sample consists of the 2001, 2002, and 2003 packing seasons. We exclude two observations during which the air-quality alerts occurred as a result of the Biscuit Fire.

Figure 6C

Panel A. The linear effect of $PM_{2.5}$ by quantile



Panel B. The linear effect of ozone by quantile



Note: This figure presents the quantile estimates for productivity based on a linear control for $PM_{2.5}$ (panel A) or ozone (panel B).

2. Estimates Based on the Parsimonious Specification

Table 2P. The Relationship between $PM_{2.5}$ and Labor Supply

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Working that day				Hours			
$PM_{2.5}$	- 0.001	0.001			- 0.017	- 0.107		
	(0.001)	(0.003)			(0.041)	(0.070)		
	[0.241]	[0.796]			[0.686]	[0.128]		
$PM_{2.5}$ 10–15			0.019	0.017			- 0.035	- 0.035
			(0.021)	(0.019)			(0.365)	(0.365)
			[0.382]	[0.384]			[0.923]	[0.923]
$PM_{2.5}$ 15–20			0.076	0.067			1.189	1.189
			(0.046)	(0.041)			(0.396)	(0.396)
			[0.100]	[0.101]			[0.003]	[0.003]
$PM_{2.5}$ 20–25			- 0.077	- 0.068			- 3.874	- 3.874
			(0.061)	(0.054)			(0.414)	(0.414)
			[0.207]	[0.205]			[0.000]	[0.000]
$PM_{2.5}$ >25			- 0.056				1.150	
			(0.039)				(0.766)	
			[0.150]				[0.133]	
R^2					0.565	0.577	0.666	0.658
N	1,275	1,238	1,275	1,238	1,162	1,143	1,162	1,143

Standard errors based on estimates clustered by date in parentheses for columns 1 through 4, standard errors based on estimates clustered by date and worker in parentheses for columns 5 through 8; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we observe $PM_{2.5}$. Columns 1 through 4 present marginal effects based on a logit model, and columns 5 through 8 present results from ordinary least squares regressions. All regressions include wind speed, a wind direction dummy variable, dew point, a rain dummy variable, an indicator variable for weekends, year dummy variables, and month dummy variables.

Table 3P. The Relationship between $PM_{2.5}$ and Productivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable:	Productivity				Logarithm of Productivity			
$PM_{2.5}$	- 0.059	- 0.049			- 0.006	- 0.007		
	(0.034)	(0.069)			(0.005)	(0.011)		
	[0.093]	[0.476]			[0.195]	[0.548]		
$PM_{2.5}$ 10–15			- 0.228	- 0.228			- 0.017	- 0.017
			(0.553)	(0.553)			(0.087)	(0.087)
			[0.681]	[0.681]			[0.847]	[0.847]
$PM_{2.5}$ 15–20			- 0.820	- 0.820			- 0.211	- 0.211
			(1.397)	(1.398)			(0.219)	(0.219)
			[0.557]	[0.557]			[0.335]	[0.335]
$PM_{2.5}$ 20–25			- 1.944	- 1.944			- 0.299	- 0.299
			(0.792)	(0.793)			(0.125)	(0.125)
			[0.014]	[0.014]			[0.016]	[0.017]
$PM_{2.5}$ >25			- 1.869				- 0.163	
			(1.084)				(0.164)	
			[0.085]				[0.320]	
R^2	0.212	0.211	0.218	0.218	0.181	0.184	0.187	0.189
N	1,162	1,143	1,162	1,143	1,162	1,143	1,162	1,143

Standard errors based on estimates clustered by date and worker in parentheses; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season, but restricted solely to the days on which we observe $PM_{2.5}$. All columns present results from ordinary least squares regressions. All regressions include wind speed, a wind direction dummy variable, dew point, a rain dummy variable, a dummy variable to indicate weekends, year dummy variables, and month dummy variables. Productivity is measured as earnings per hour.

Table 4P. Robustness Checks

	Baseline estimates (1)	Exclude meterological controls (2)	Control flexibly for temperature (3)	Control for additional pollutants (4)	Add worker fixed effects (5)
$PM_{2.5}$	- 0.059	- 0.037	- 0.010	- 0.049	- 0.059
	(0.032)	(0.032)	(0.039)	(0.026)	(0.032)
	[0.067]	[0.254]	[0.792]	[0.065]	[0.067]
R^2	0.212	0.196	0.237	0.217	0.212
N	1,162	1,162	1,162	1,162	1,162

	Aggregate to six-day PM-measurement periods (6)	Median regression (7)	Minimum wage binds (8)	Censored median regression (9)	Low-quality packing (10)
$PM_{2.5}$	(Not applicable.)	- 0.046	0.001	- 0.069	- 0.002
		(0.041)	(0.006)	(0.044)	(0.002)
		[0.262]	[0.918]	[0.115]	[0.277]
R^2					0.139
N		1,162	1,162	989	984

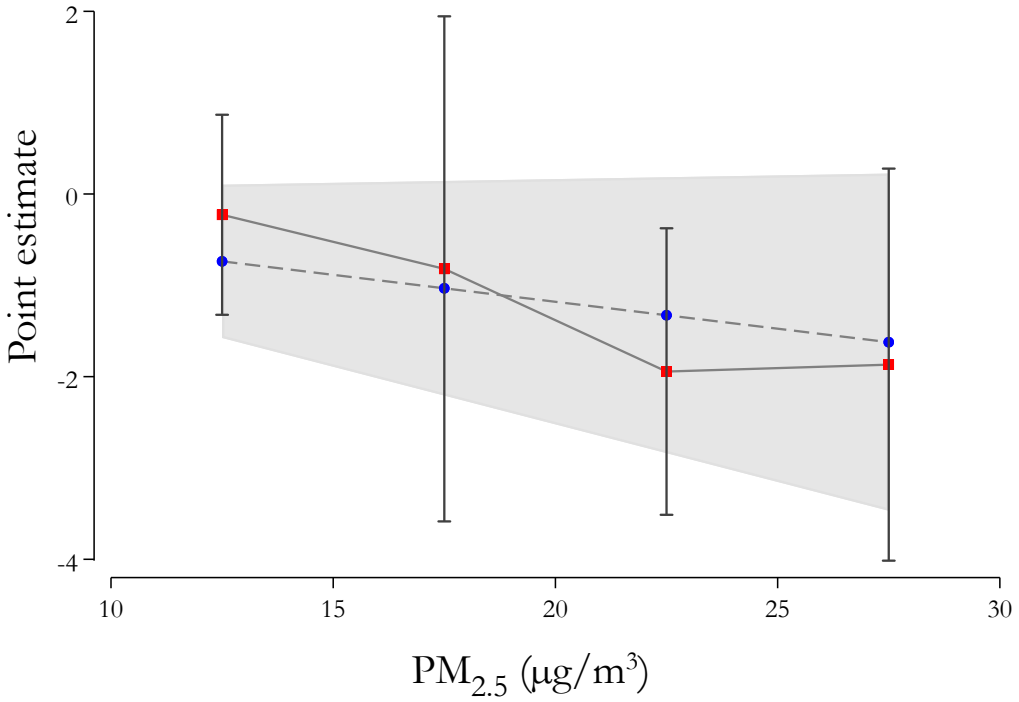
Standard errors based on estimates clustered by date and worker in parentheses for columns 1 through 5, standard errors based on estimates clustered by date in parentheses for columns 6 through 10; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we observe $PM_{2.5}$. All regressions include a dummy variable to indicate weekends, year dummy variables, and month dummy variables. All regressions except column 2 include wind speed, a wind direction dummy variable, dew point, and a rain dummy variable. Column 3 controls for temperature flexibly by including a series of indicator variables for each 5°F. Column 4 includes nitrogen dioxide, carbon monoxide, and coarse PM. In all regressions except for columns 8 and 10, the dependent variable is productivity during the regular-time shift, which is measured in earnings per hour. Column 8 uses whether the minimum wage binds as the dependent variable and column 10 uses "low-quality packing" as the dependent variable. Column 8 presents marginal effects from a logit model; column 10 presents estimates from a linear probability model.

Table 5P. The Relationship between $PM_{2.5}$ and Overtime Productivity

	Overtime hours (1)	Regular-time produtivity (2)	Overtime productivity (3)	Overtime productivity (4)
$PM_{2.5}$	- 0.036	0.255	- 0.102	- 0.255
	(0.019)	(0.024)	(0.017)	(0.056)
	[0.060]	[0.000]	[0.000]	[0.000]
R^2	0.283	0.295	0.295	0.413
N	1,162	332	332	332

Standard errors based on estimates clustered by date and worker in parentheses; associated p-values in brackets. The sample consists of worker-day observations over the 2001, 2002, and 2003 pear-packing season for which we observe $PM_{2.5}$. All regressions include wind speed, a wind direction dummy variable, dew point, a rain dummy variable, a dummy variable to indicate weekend, year dummy variables, and month dummy variables. In columns 1 and 2, the dependent variable is the number of overtime hours worked. The dependent variable in column 3 is regular-time productivity and in columns 4 and 5 is overtime earnings, both limited to the sample of worker-days for which overtime hours exists. Productivity is measured in earnings per hour, though overtime productivity is deflated by 1.5 to account for time-and-a-half overtime pay.

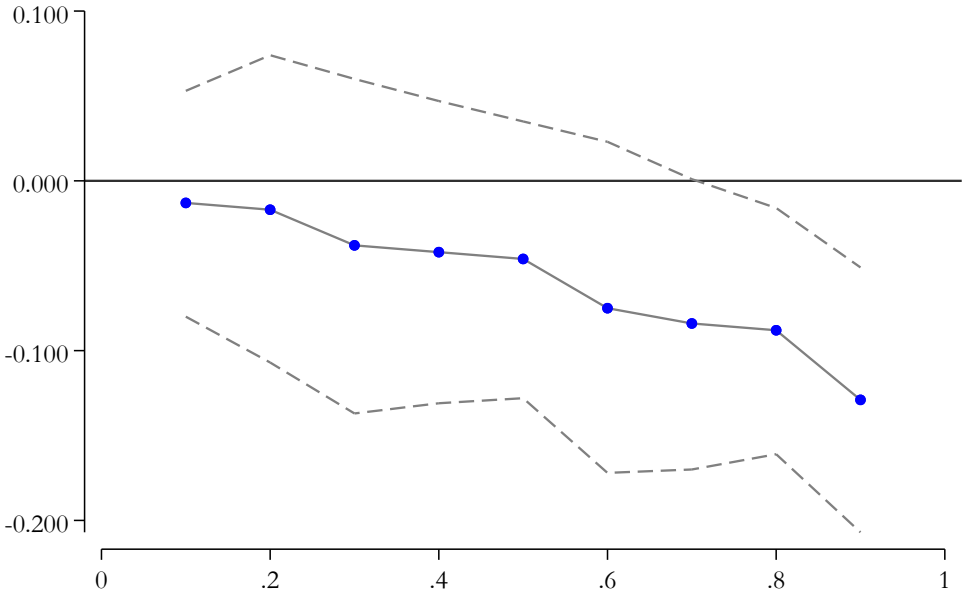
Figure 5P



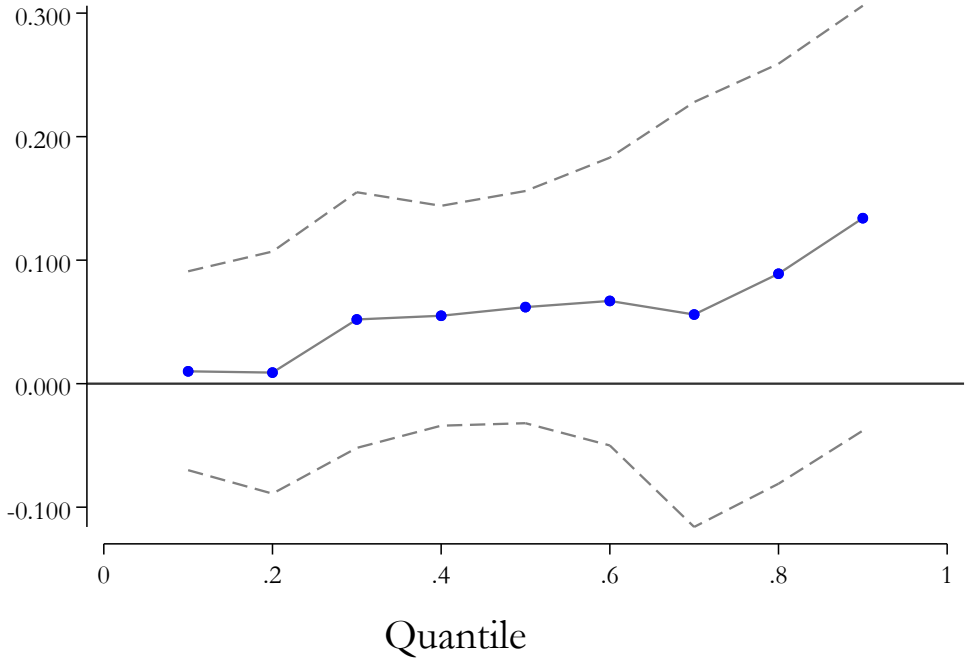
Note: This figure presents the implied effects of $PM_{2.5}$ on productivity based on estimates reported in Table 3P, columns 1 (linear) and 3 (nonlinear).

Figure 6P

Panel A. The linear effect of $PM_{2.5}$ by quantile



Panel B. The linear effect of ozone by quantile



Note: This figure presents the quantile estimates for productivity based on a linear control for $PM_{2.5}$ (panel A) or ozone (panel B).