Recovery from Disasters

Data Issues and Regressions

December 14, 2021

Here, in Table 1, I document main issues or limitations of data and suggest how to deal with them. As you can see, there are very standard problems in this panel, such as missing data and unbalancedness. To illustrate whether results are robust to applying "possible solutions" [1] to these issues, I create plots of coefficients. These plots contain 3 groups of models: models in levels, growth rate models, and growth rate models, where I add a constant to 0 values in before log transforming dependent variables.

Just to recap, this is the main level regressions:

$$y_{it} = \sum_{l=0}^{L} \beta_l S_{i,t-l} + \alpha_i + \gamma_t + \varepsilon_{it},$$

where y_{it} is a particular variable of a firm i in year t, $S_{i,t-l}$ is one of the shaking measures, lagged L times. α_i is a plant fixed effect, γ_t is a year time fixed effect. ε_{it} is a standard error, clustered on a firm-level. S in this document is $mpga\_aw$[2]. For growth rate models with a constant, I just add 0.001 to all dependent variables.

This is the basic growth rate model:

$$\Delta ln(y_{it}) = \sum_{l=0}^{L} \beta_l S_{i,t-l} + \alpha_i + \gamma_t + \varepsilon_{it},$$

where $\Delta ln(y_{it}) \equiv ln(y_{it}) - ln(y_{it-1})$ is the growth rate of a particular variable of a firm i in year t, and all other parts are identical to the specification in levels.

First, I want to check consistency between results for 5 and 10 lags. Second, I run regressions for a subsample where all observations have all 10 lags (issue III). Third, I run regressions only for firms that have no missing values (different subsample for each variable).

[1] It's a column in Table 1

[2] average area-weighted of maximum of PGA/PGV (peak ground acceleration/velocity) over each gridcell in each year

Table 1: List of data problems (to be continued).

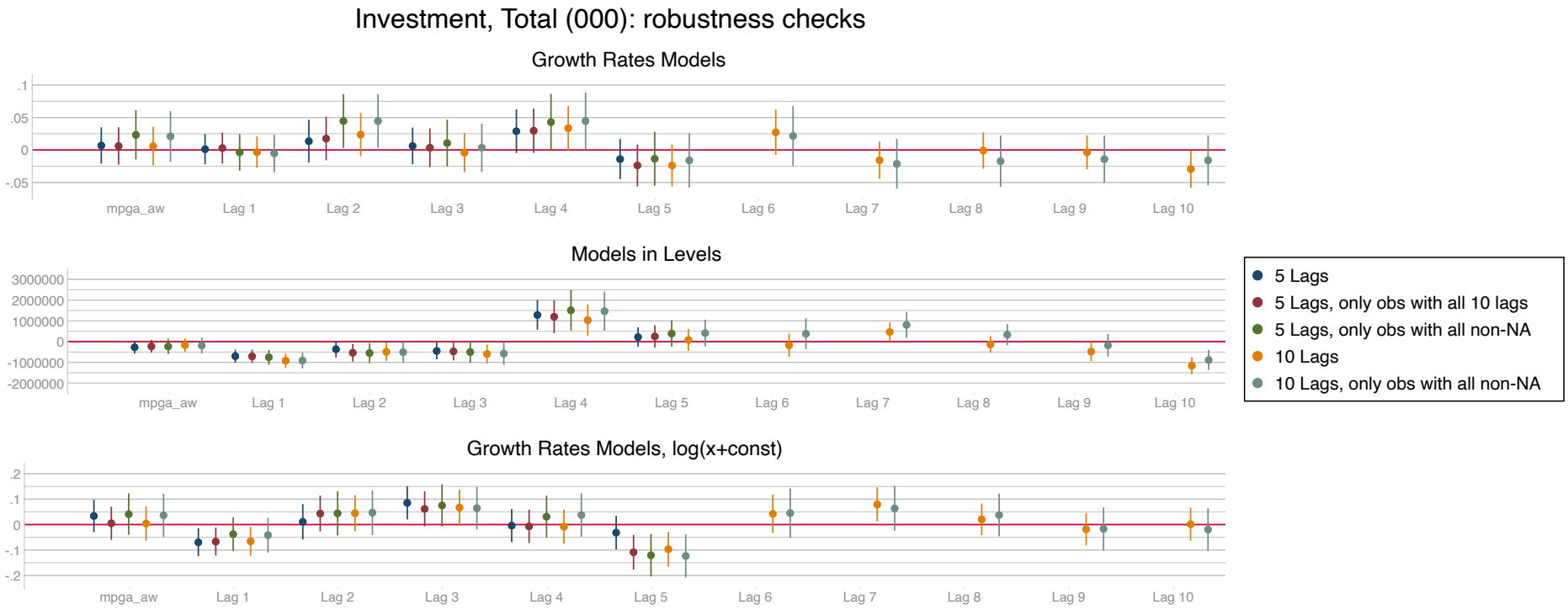
Issue	Possible solutions	Did I implement?
I. 0 values in firm data - issue with log specifications	1. Adding a constant (not recommended, but I do it for illustration purposes)	+
	2. Dropping 0 observations (creates missing values)	+
	3. Imputing from data	-
	4. Using alternatives to log transformation	-
II. Missing values in firm data	1. Only keeping firms that have no missing values	+
	2. Imputing from data	-
III. Missing values in EQ data (pre-1973)	1. Keep only observations with N lags (to compare results from N and N-1, N-2... lags)	+
IV. Variation in a number of observations for each firm (entrances/exits/problems with reporting?)	?	
V. Variation in reporting years for each variable (e.g. capital in Indonesia only since 1988, all data - since 1975)	?	
VI. Little variation in number of EQ (since it's very rare) (cannot obtain estimates every time)	1. Using shaking measures instead	+

I also have all the results in appendix for each country for main variables.

I admit that before I plotted those, I thought that results were more sensitive to changes of subsamples. It turns out to be not completely true: results may or may not be statistically significant, but they often have the same sign and similar p-values. Look at Figure 1. Results for growth rates model for lag 2 depend on whether we use all the data or only non-missing sub-sample. Results from a model where I added a constant before log transformation seem more noizy and have less consistent pattern. For this and the rest of the graphs, results are mostly consistent among robustness checks but inconsistent among specifications.

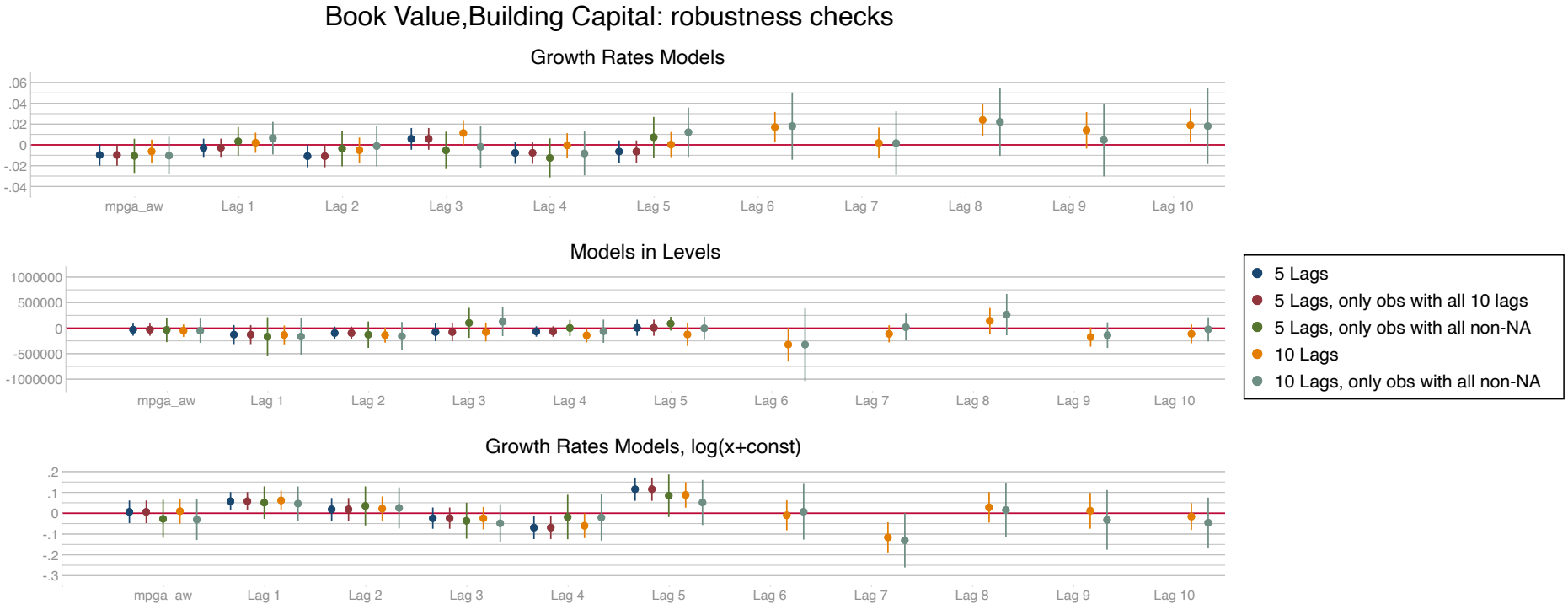
Indonesia. Figure 1.

Figure 1

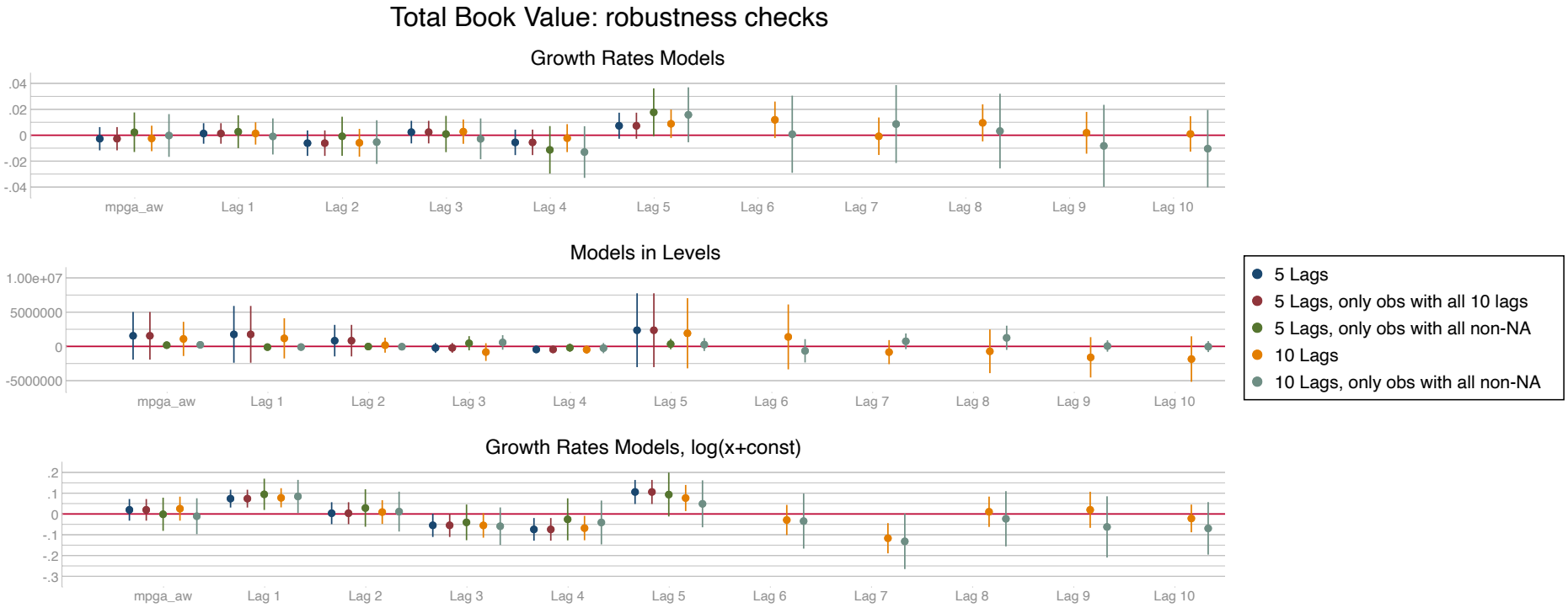


Appendices start here.

Appendix. Indonesia



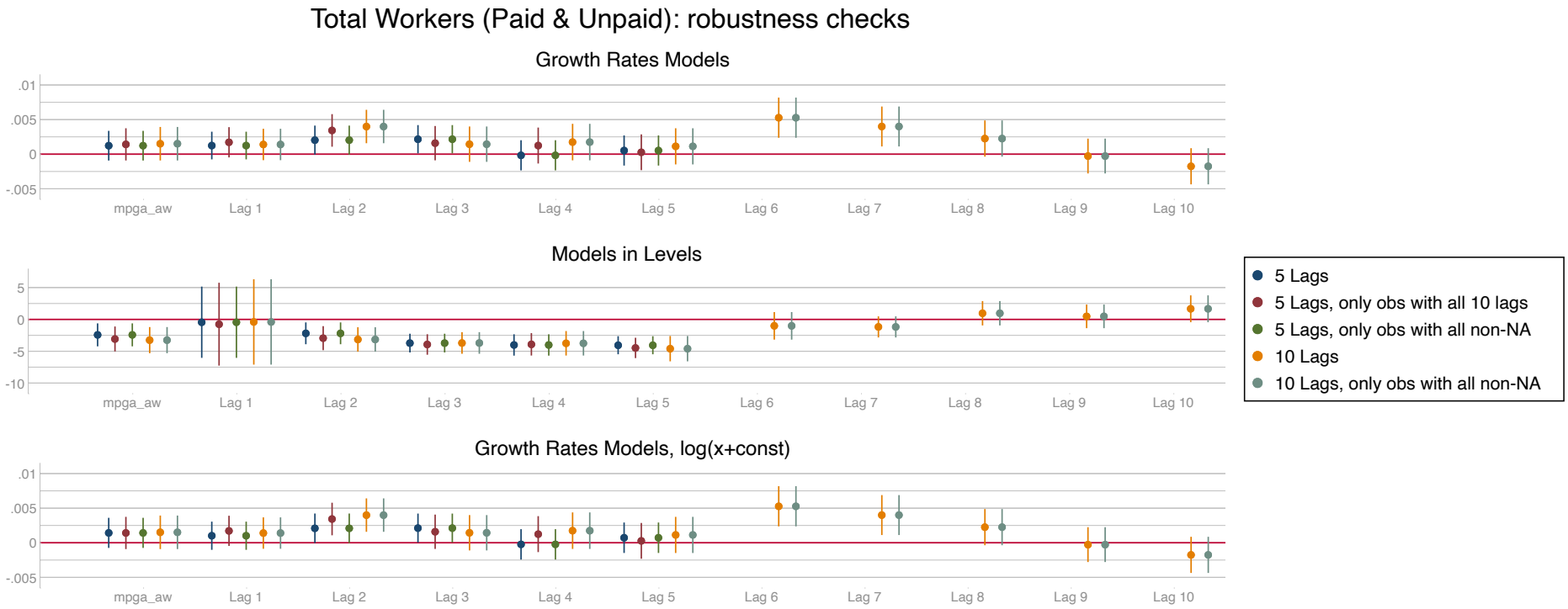
Indonesia


Total Book Value: robustness checks
Growth Rates Models
.04
.02
0
-.02
-.04
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1.00e+07
5000000
0
-5000000
Growth Rates Models, log(x+const)
.2
.1
-.1
-.2
-.3
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

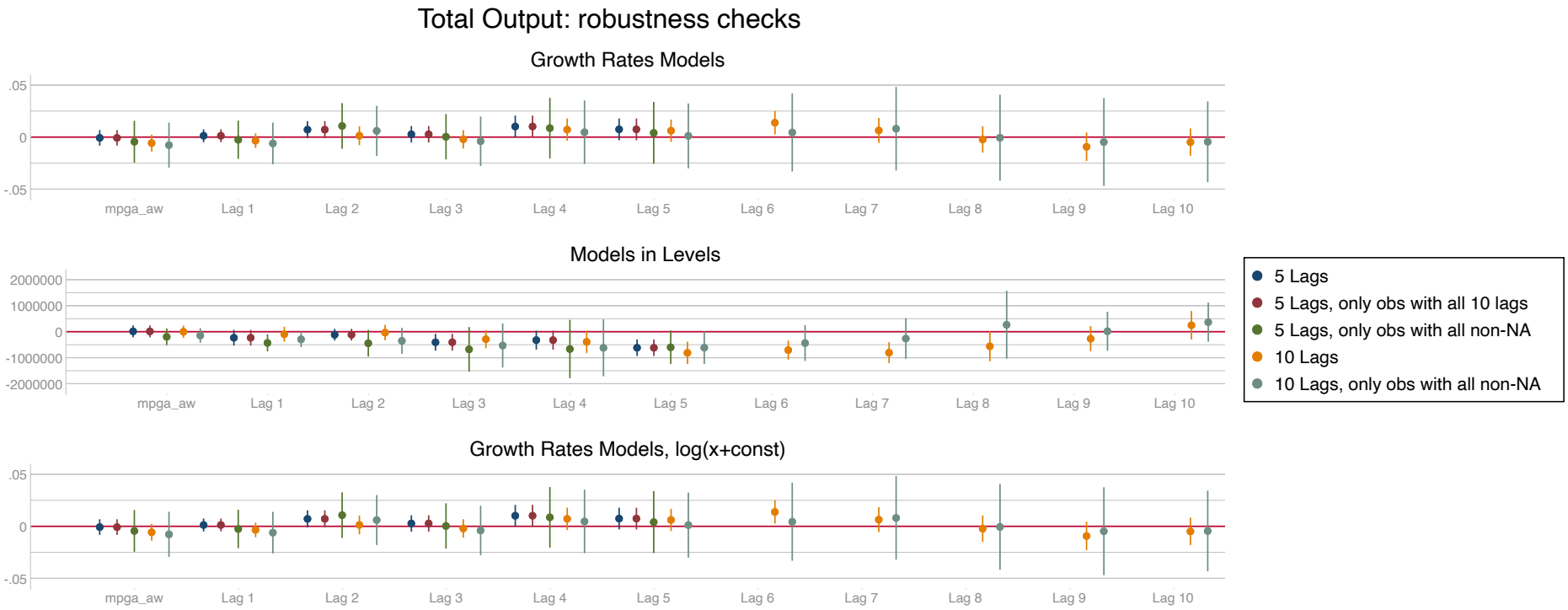
Indonesia


Goods Exported: robustness checks
Growth Rates Models
.1
0
-.1
-.2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
200000
0
-200000
-400000
-600000
-800000
Growth Rates Models, log(x+const)
.4
.2
0
-.2
-.4
-.6
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

Indonesia


Total Workers (Paid & Unpaid): robustness checks
Growth Rates Models
.01
.005
0
-.005
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
5
0
-5
-10
Growth Rates Models, log(x+const)
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

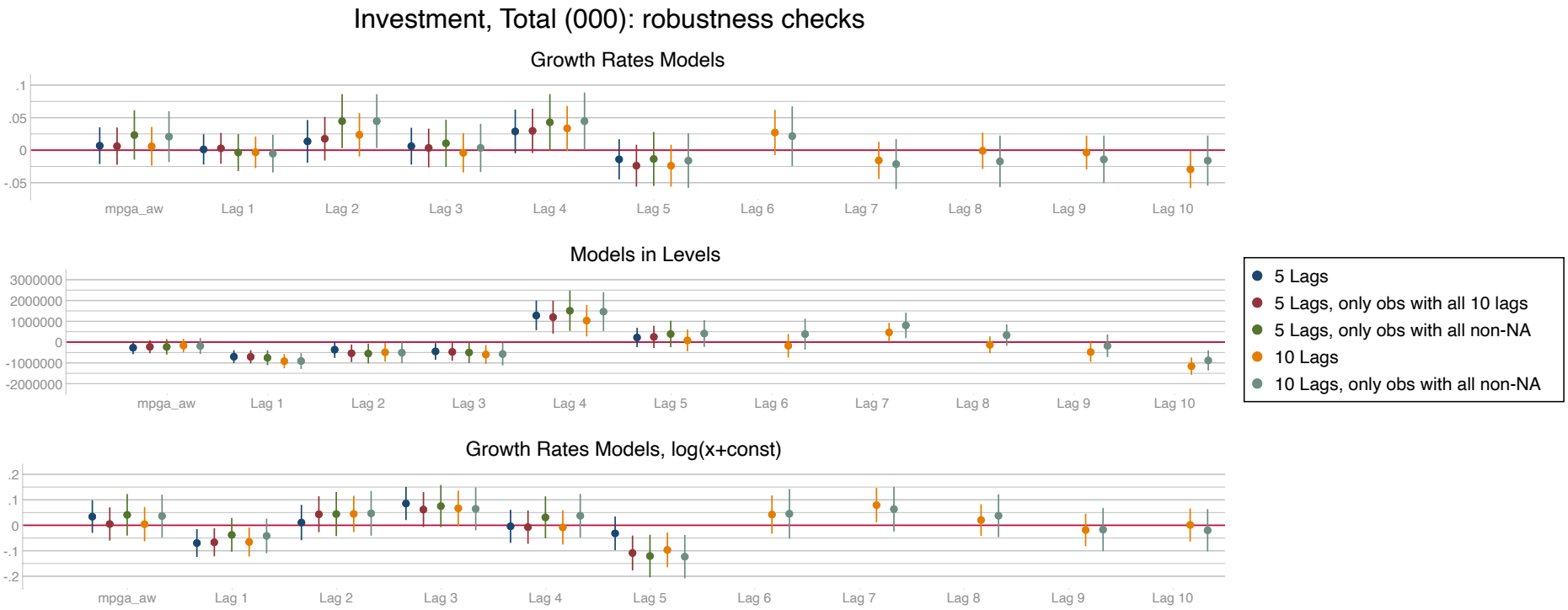
Indonesia


Total Output: robustness checks
Growth Rates Models
.05
0
-.05
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
2000000
1000000
0
-1000000
-2000000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
.05
0
-.05
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

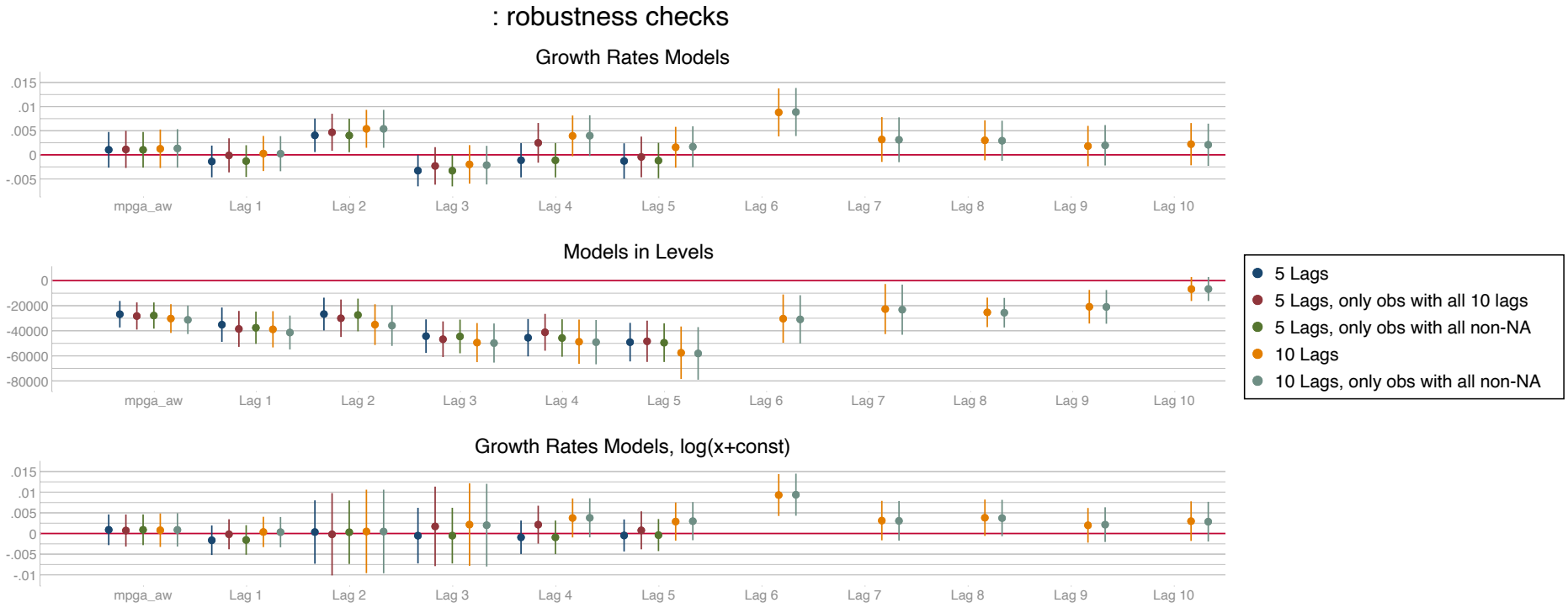
Indonesia


Overall Stock (raw, semi, finished), end of year (000): robustness checks
Growth Rates Models
.4
.2
0
-.2
-.4
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1000000
0
-1000000
-2000000
Growth Rates Models, log(x+const)
1
0
-1
-2
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

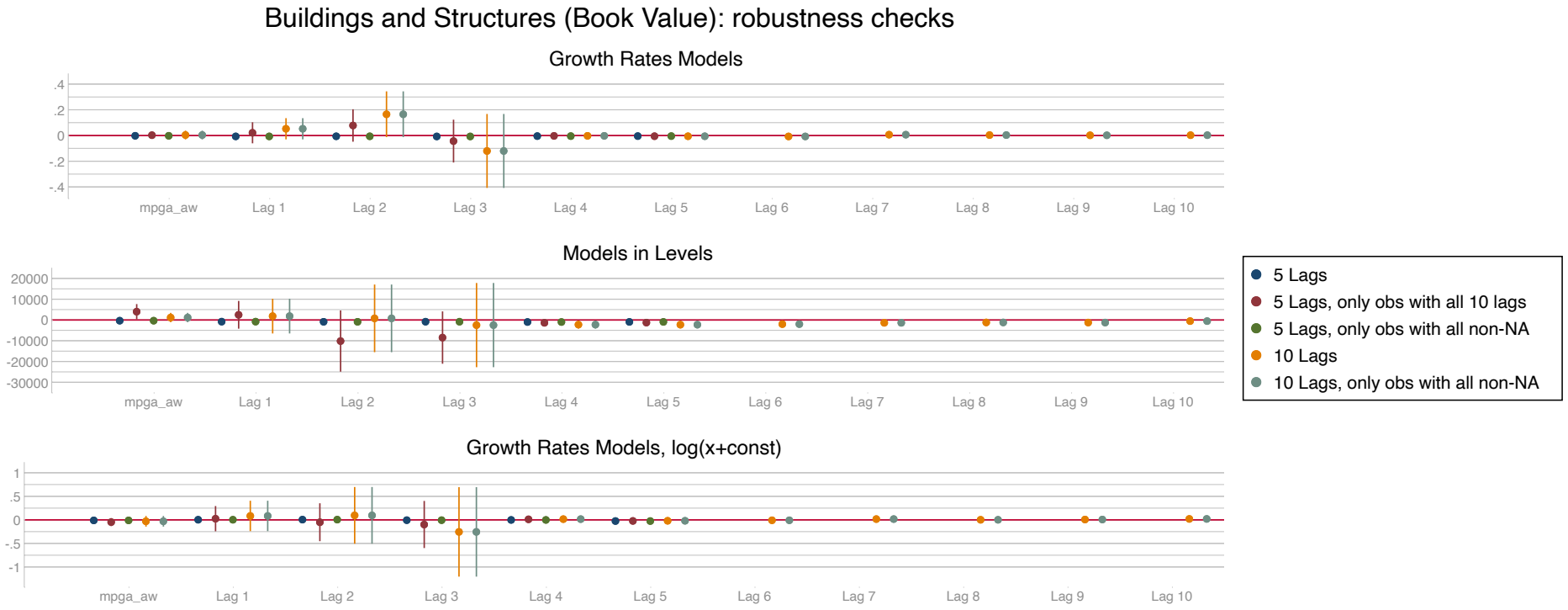
Indonesia


Investment, Total (000): robustness checks
Growth Rates Models
.1
.05
0
-.05
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
3000000
2000000
1000000
0
-1000000
-2000000
Growth Rates Models, log(x+const)
.2
.1
0
-.1
-.2
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

Indonesia: wages



Appendix. Colombia



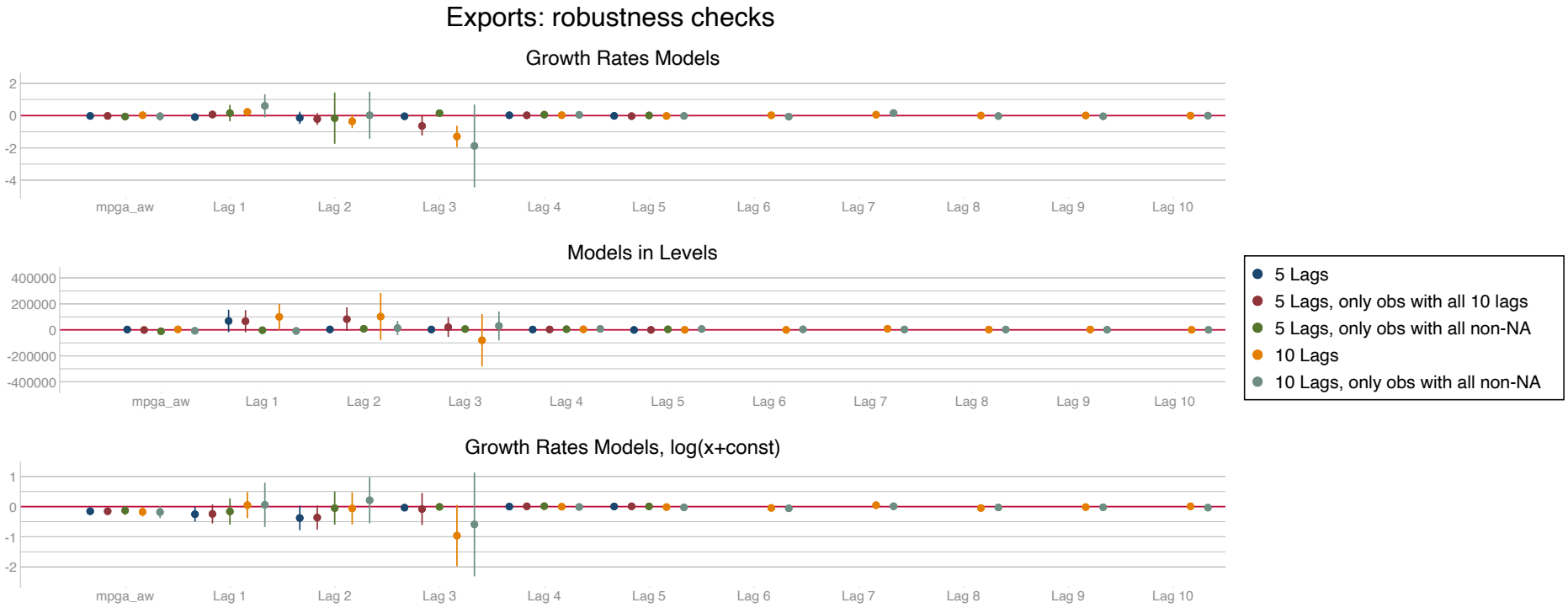
Colombia


Total Assets Book Value: robustness checks
Growth Rates Models
.2
.1
0
-.1
-.2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
100000
0
-100000
-200000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
.4
.2
0
-.2
-.4
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

Colombia


Total Domestic Sales: robustness checks
Growth Rates Models
.5
0
-.5
-1
-1.5
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1000000
500000
0
-500000
-1000000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
2
1
0
-1
-2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

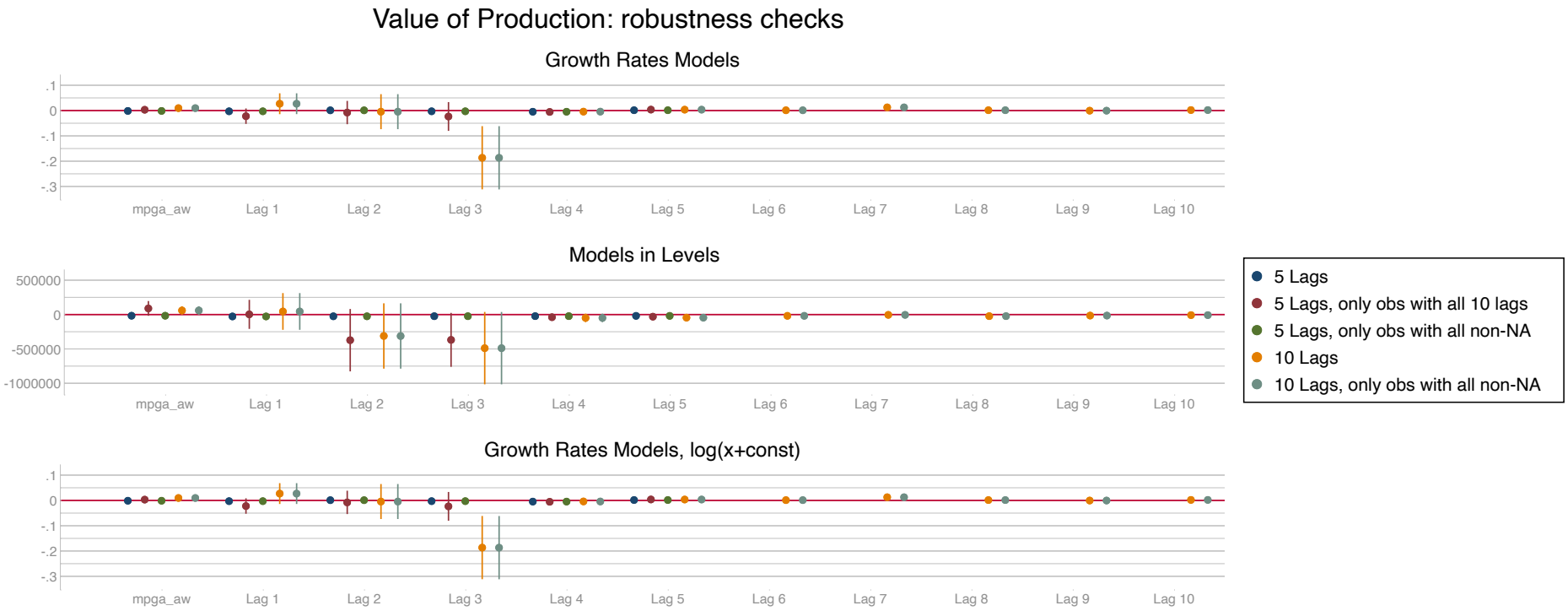
Colombia


Exports: robustness checks
Growth Rates Models
2
0
-2
-4
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
400000
200000
0
-200000
-400000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
1
0
-1
-2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

Colombia


Total Employment: robustness checks
Growth Rates Models
.1
0
-.1
-.2
-.3
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
20
10
0
-10
-20
Growth Rates Models, log(x+const)
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

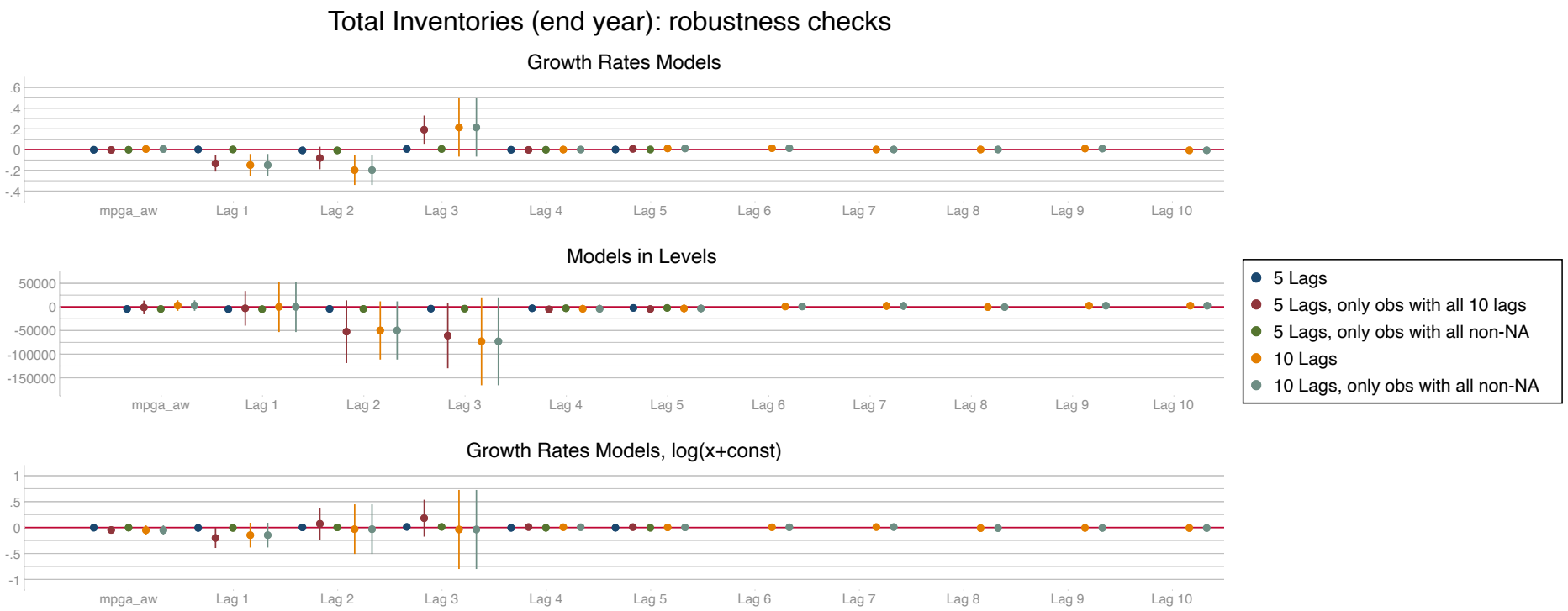
Colombia


Value of Production: robustness checks
Growth Rates Models
.1
0
-.1
-.2
-.3
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
500000
0
-500000
-1000000
Growth Rates Models, log(x+const)
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

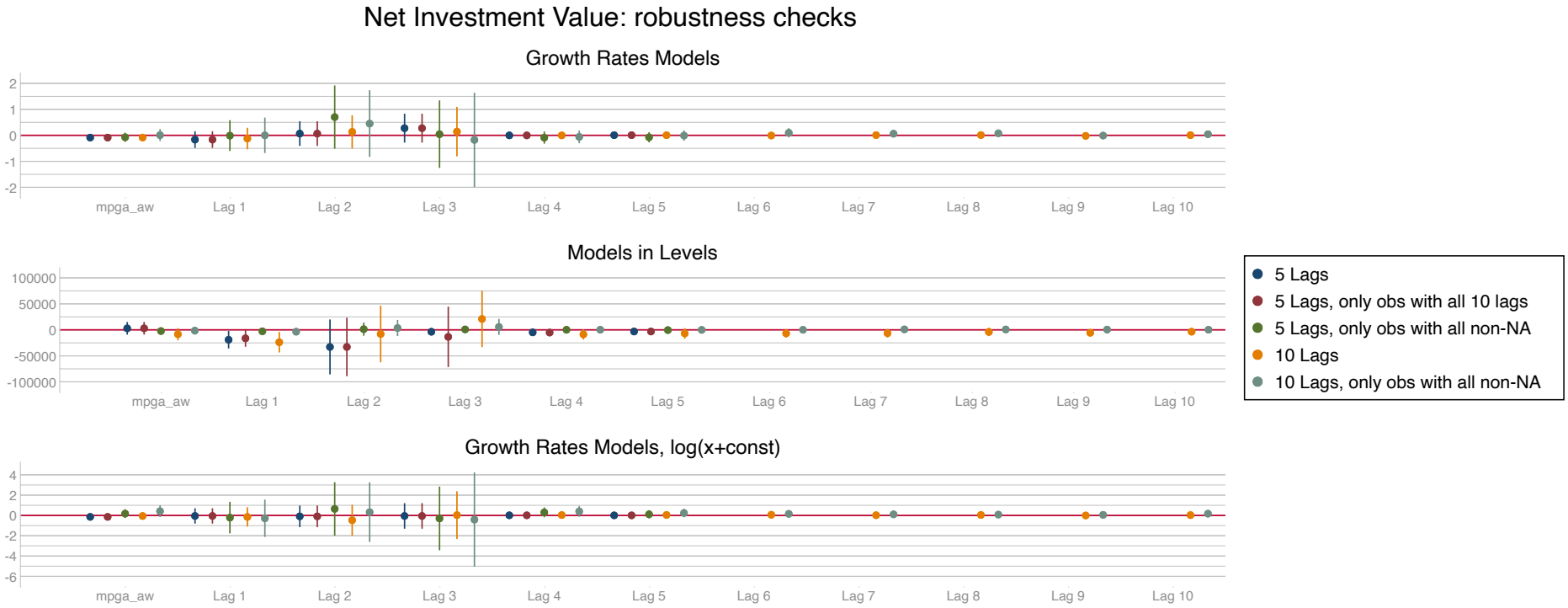
Colombia


Total Sales: robustness checks
Growth Rates Models
.1
0
-.1
-.2
-.3
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
500000
0
-500000
-1000000
Growth Rates Models, log(x+const)
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

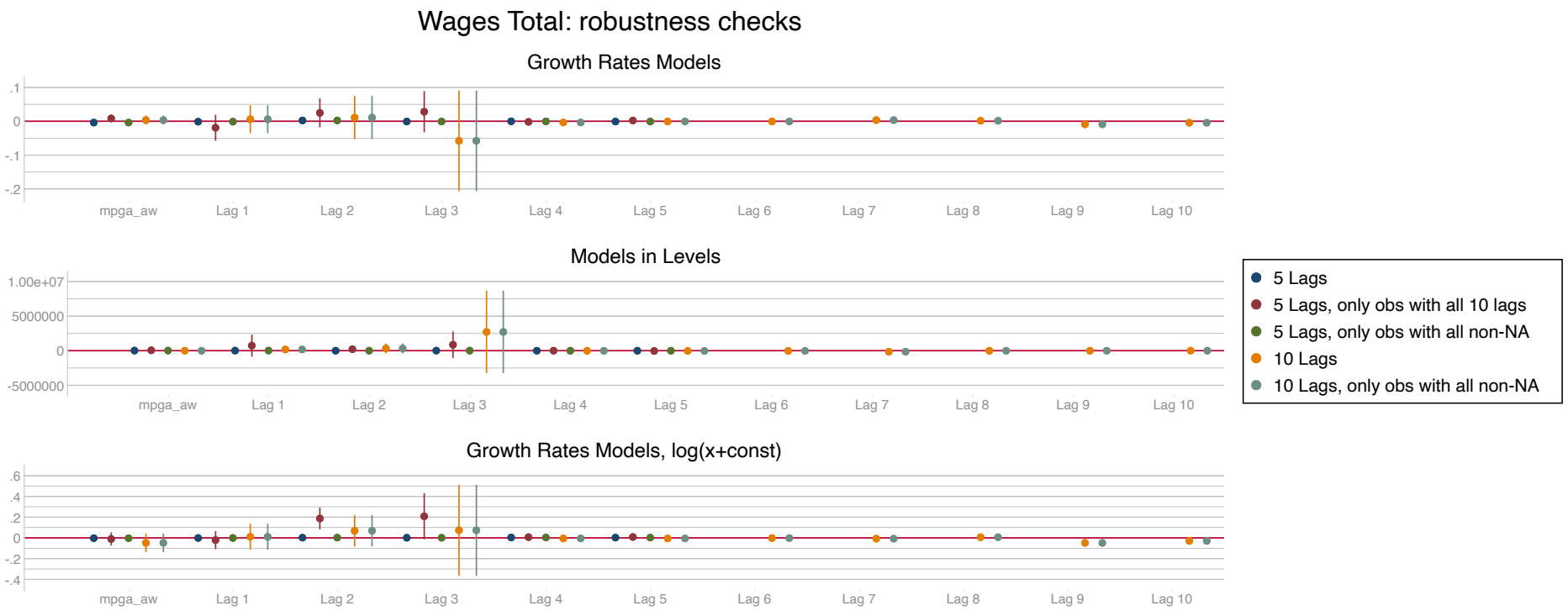
Colombia


Total Inventories (end year): robustness checks
Growth Rates Models
.6
.4
.2
0
-.2
-.4
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
50000
0
-50000
-100000
-150000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
1
.5
0
-.5
-1
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

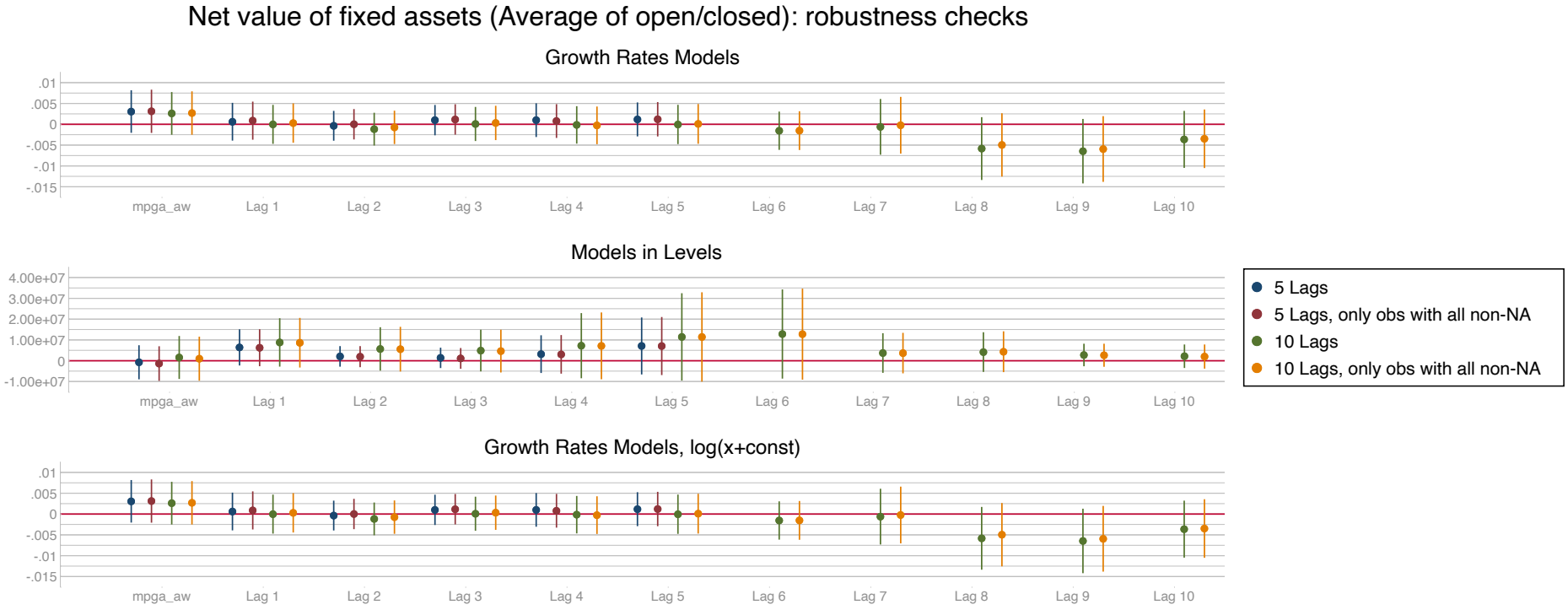
Colombia


Net Investment Value: robustness checks
Growth Rates Models
2
1
0
-1
-2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
100000
50000
0
-50000
-100000
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
4
2
0
-2
-4
-6
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

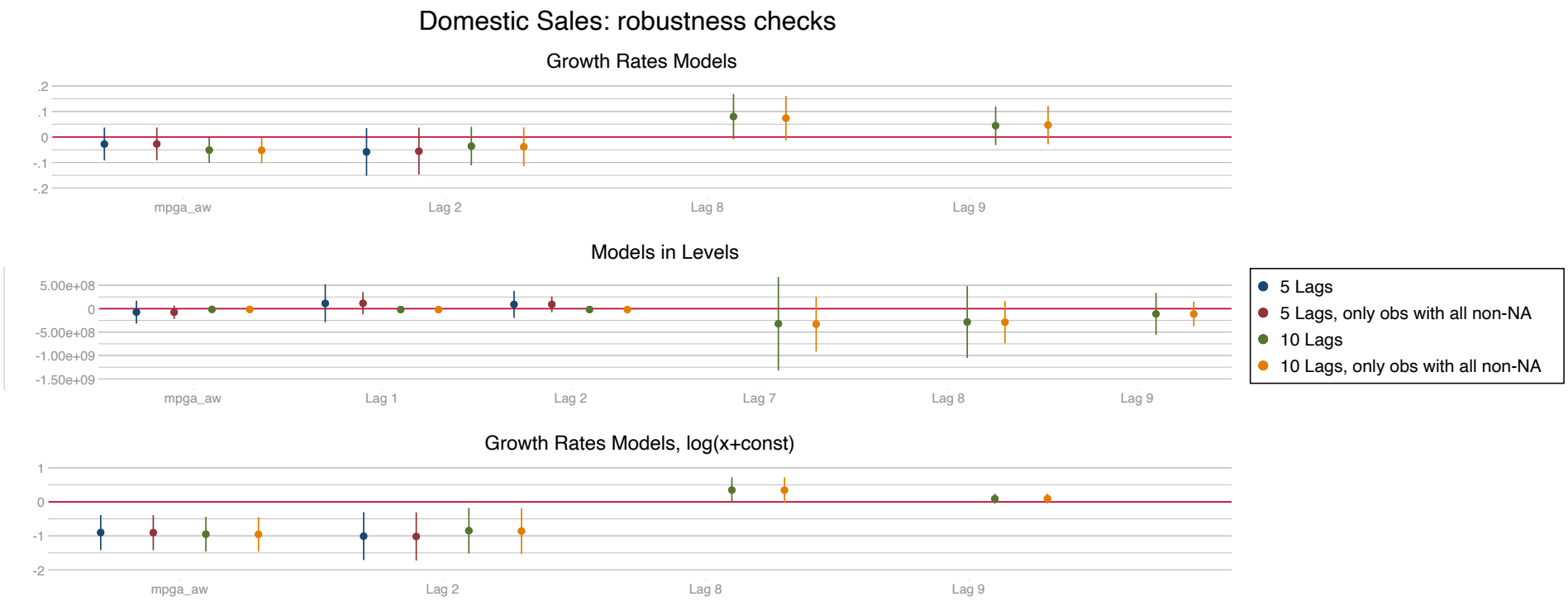
Colombia


Wages Total: robustness checks
Growth Rates Models
.1
0
-.1
-.2
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1.00e+07
5000000
0
-5000000
Growth Rates Models, log(x+const)
.6
.4
.2
0
-.2
-.4
5 Lags
5 Lags, only obs with all 10 lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

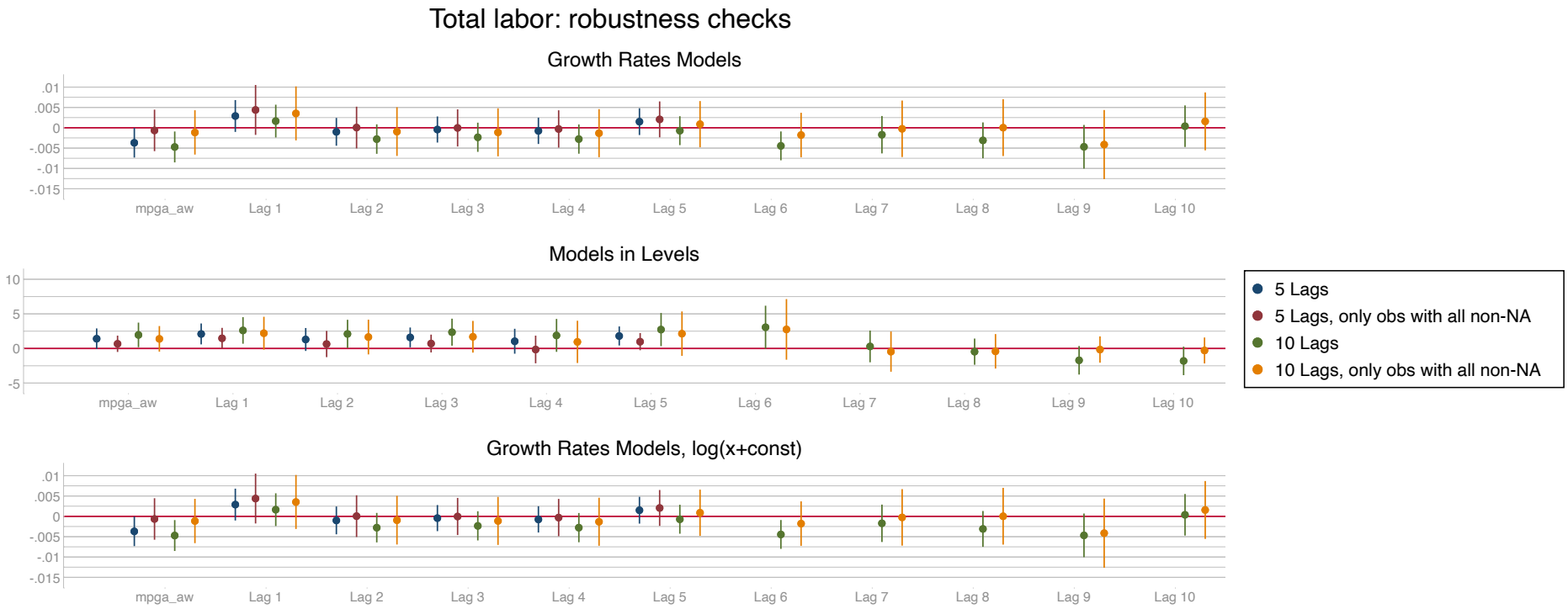
Appendix. India



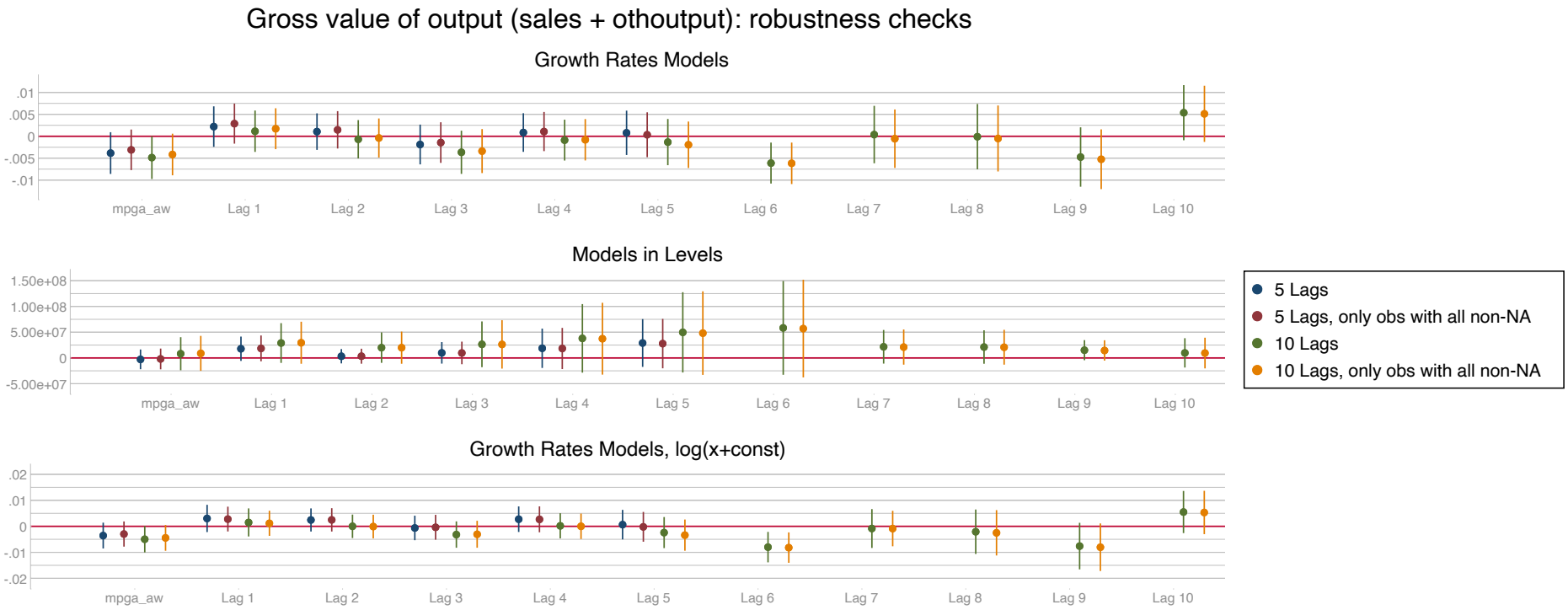
India


Domestic Sales: robustness checks
Growth Rates Models
.2
.1
0
-.1
-.2
mpga_aw
Lag 2
Lag 8
Lag 9
Models in Levels
5.00e+08
0
-5.00e+08
-1.00e+09
-1.50e+09
mpga_aw
Lag 1
Lag 2
Lag 7
Lag 8
Lag 9
Growth Rates Models, log(x+const)
1
0
-1
-2
mpga_aw
Lag 2
Lag 8
Lag 9
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

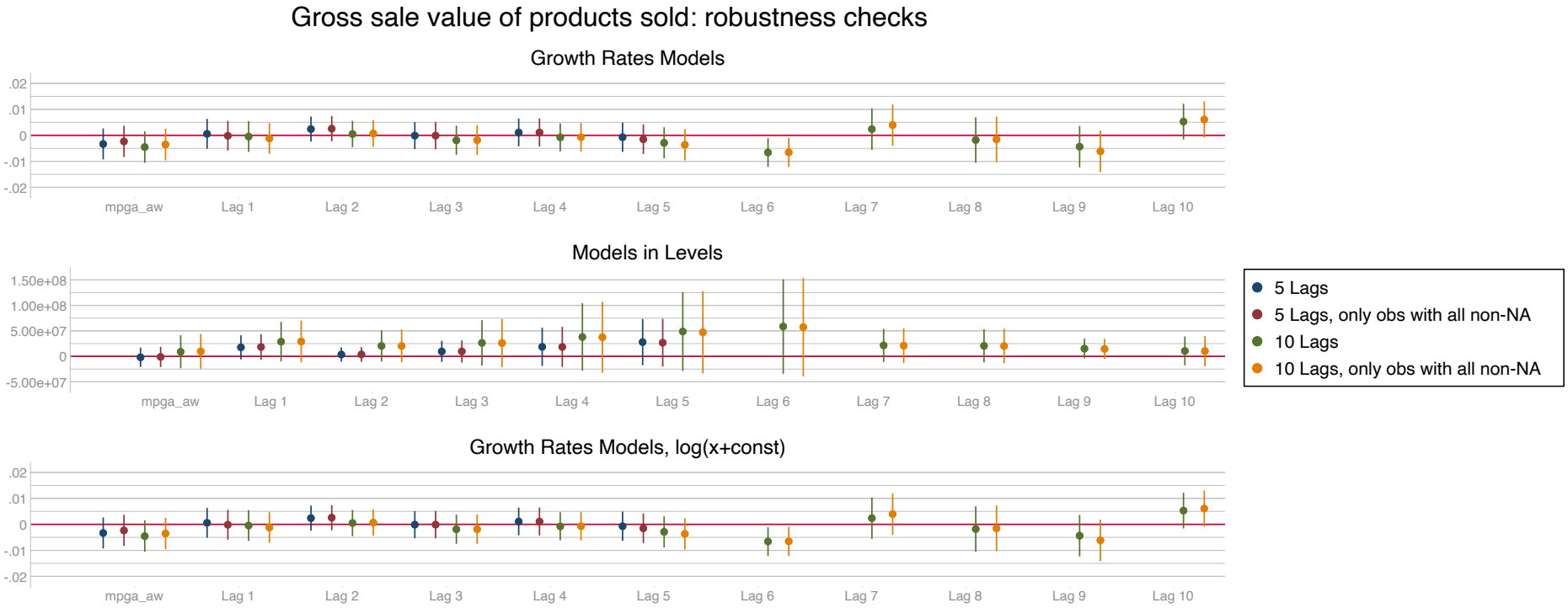
India


Total labor: robustness checks
Growth Rates Models
.01
.005
0
-.005
-.01
-.015
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
10
5
0
-5
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
.01
.005
0
-.005
-.01
-.015
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

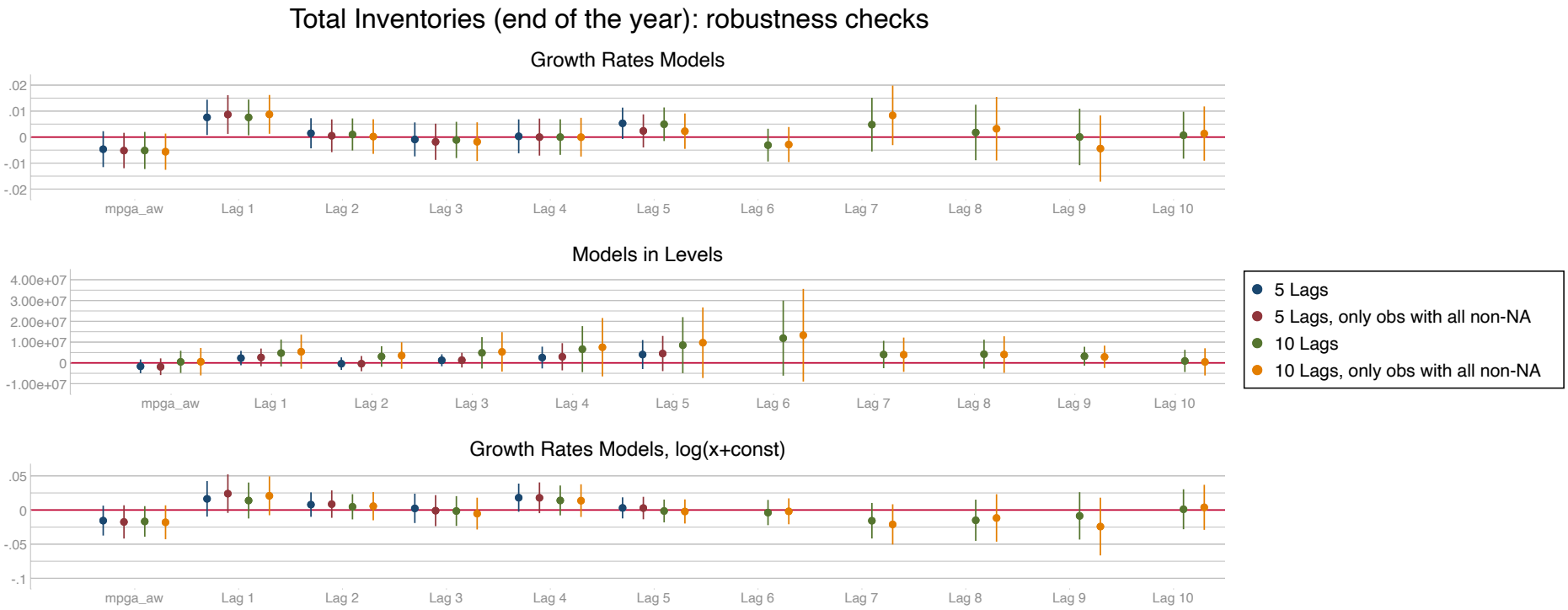
India


Gross value of output (sales + othoutput): robustness checks
Growth Rates Models
.01
.005
0
-.005
-.01
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1.50e+08
1.00e+08
5.00e+07
0
-5.00e+07
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
.02
.01
0
-.01
-.02
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

India


Gross sale value of products sold: robustness checks
Growth Rates Models
.02
.01
0
-.01
-.02
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1.50e+08
1.00e+08
5.00e+07
0
-5.00e+07
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Growth Rates Models, log(x+const)
.02
.01
0
-.01
-.02
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

India


Total Inventories (end of the year): robustness checks
Growth Rates Models
.02
.01
0
-.01
-.02
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
4.00e+07
3.00e+07
2.00e+07
1.00e+07
0
-1.00e+07
Growth Rates Models, log(x+const)
.05
0
-.05
-.1
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA

India


Total wages & salaries: robustness checks
Growth Rates Models
.01
.005
0
-.005
-.01
mpga_aw
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Models in Levels
1000000
500000
0
-500000
Growth Rates Models, log(x+const)
5 Lags
5 Lags, only obs with all non-NA
10 Lags
10 Lags, only obs with all non-NA