# Estimating the Economic Costs of Organized Crime By Generalized Synthetic Control Methods

Martin Becker
Statistics and Econometrics
Saarland University

Stefan Klößner*
Statistics and Econometrics
Saarland University

August 29, 2016

**Abstract**

The economic costs of organized crime have been estimated for the case of Southern Italy by Pinotti (2015) who finds a loss in GDP per capita by 16 % due to the advent of Italian mafia in the regions Apulia and Basilicata. We replicate the results of Pinotti (2015), derived using the synthetic control method, both in a narrow and wider sense: first, using different software, we find slightly lower estimates for the loss in GDP per capita, while using more general synthetic control methods leads to results very similar to those of Pinotti (2015).

*Correspondence to: Stefan Klößner, Saarland University, Campus C3 1, 66123 Saarbrücken, Germany. E-mail: S.Kloessner@mx.uni-saarland.de

# 1 Introduction

In a recent paper, Pinotti (2015) analyzes the economic costs of organized crime by considering evidence from Southern Italy. In particular, the regions Apulia and Basilicata are investigated, to which Italian mafia expanded during the late 1970s. For estimating the effect of organized crime on the economy, Pinotti (2015) uses the synthetic control method (SCM) developed by Abadie and Gardeazabal (2003) and Abadie et al. (2010). SCM rests upon a comparison of the actual economic development of the affected regions to a counterfactual scenario which is calculated from economic data of unaffected regions. More precisely, the counterfactual scenario is constructed by determining so-called donor weights for the unaffected regions from the control group, with these weights describing how the treated region is to be synthesized. When looking for these donor weights, two goals are pursued: treated and synthetic region must resemble each other as well as possible, both with respect to the pre-treatment values of the variable of interest, GDP per capita, and with respect to so-called economic predictors: the latter are the initial level of GDP per capita, investment rate, human capital, population density, and the shares of value added of agriculture, industry, market services, and non-market services.[1]

In Section 2, we replicate the results of Pinotti (2015) in a narrow sense, i.e. we use exactly the same specification and exactly the same methodological approach, but with a different software package. Using the statistical software R in combination with R packages MSCMT and DEoptim,[2] we find that the results derived by Pinotti (2015) using Stata are suboptimal for the main specification as well as for some of the robustness studies. After discussing the differences between our results and those of Pinotti (2015), we turn in Section 3 to replication in a wider sense. As a further check of robustness, but also in order to investigate the channels through which organized crime affects economic activity, we modify the specification in several ways: we remove economic predictors, we treat the economic predictors as time series (SCMT approach), and we consider several dependent variables simultaneously (MSCM and MSCMT approaches).[3] After discussing the similarities and differences with respect to the results of Pinotti (2015), Section 4 concludes.

# 2 Replication in a Narrow Sense

For synthesizing Apulia and Basilicata, Pinotti (2015) employs an SCM approach using the following specification: dependent variable is GDP per capita, economic predictors are the mean of GDP per capita, investment rate, human capital, population density, and the shares of value added of agriculture, industry, market services, and non-market services. For determining the synthetic weights, the differences between these variables for the treated region and the synthetic control are minimized over the time span 1951-60.[4] Pinotti (2015) reports as optimal weights 62.4% for Abruzzo and 32.6% for Molise, derived using the commercial software Stata. Simple calculations show that these weights correspond to a root mean squared prediction error (RMSPE) with respect to GDP per

[1]For more details and formulas, see (Pinotti, 2015, F214ff.).

[2]See R Core Team (2016), Becker and Klößner (2016b), Ardia et al. (2015).

[3]SCMT, MSCM, MSCMT have been introduced in Klößner and Pfeifer (2016).

[4]Notice, however, that except for GDP per capita and population density, all variables are available since 1960 only.

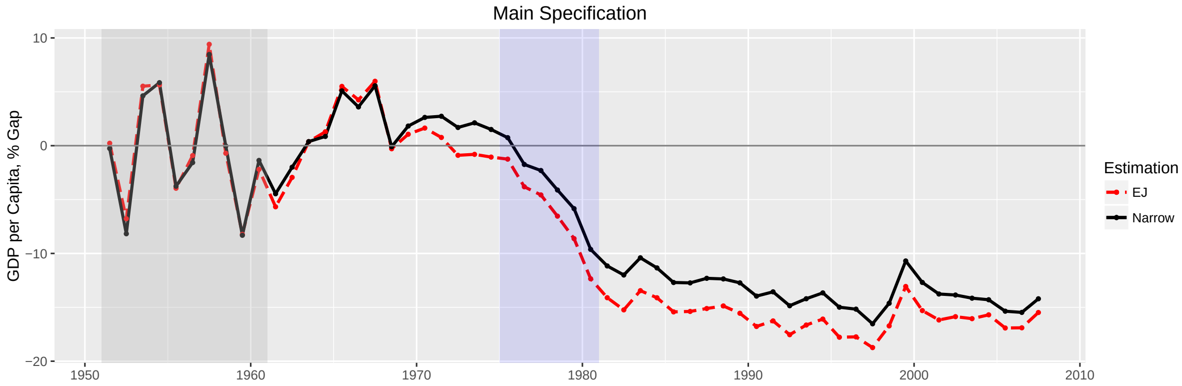


Figure 1: GDP gaps for replication in a narrow sense

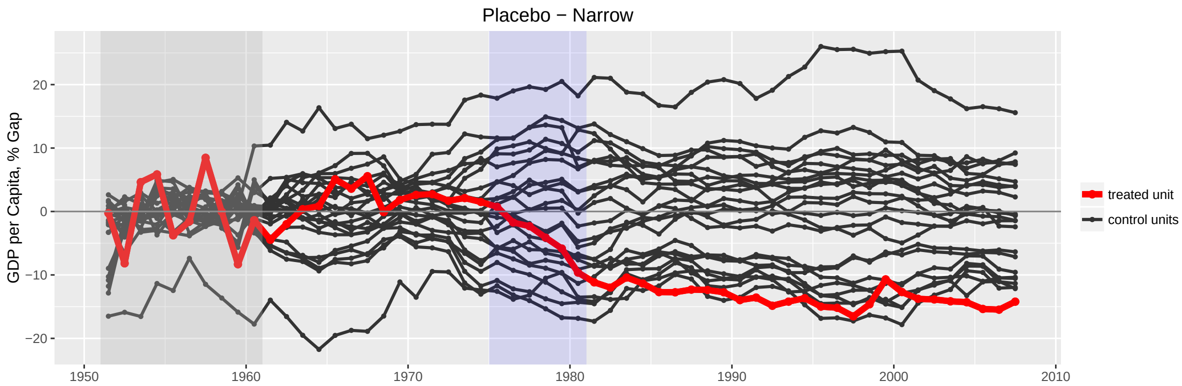


Figure 2: Placebo plot for replication in a narrow sense

capita over 1951-60 of 130.14.[5] For replicating the results of Pinotti (2015), we use the publicly available statistical software R in combination with packages MSCMT and DEoptim.[6] In contrast to Pinotti (2015), we find optimal weights for the donor countries to be 43.04% for Abruzzo, 48.93% for Molise, and 8.03% for Sardinia. Compared to the suboptimal solution of Pinotti (2015), the RMSPE reduces to 128.95.[7]

Figure 1 shows the estimated effect of organized crime on GDP per capita as estimated by Pinotti (2015) (red dashed line, denoted 'EJ') and the corresponding narrow replication (black solid line, denoted 'Narrow').[8] The effects are of a similar magnitude, however, it is clearly visible that the results of the narrow replication imply a slighly smaller loss of GDP per capita than those of Pinotti (2015).

In Pinotti (2015), a so-called placebo study is conducted in order to gain information about the significance of the loss in GDP per capita that Apulia and Basilicata suffer from after the mafia expanded to these regions. In a placebo study, the roles of treated and unaffected regions are swapped and gaps for the actually treated unit are contrasted with the corresponding artificial gaps. Figure 2 depicts the results of this placebo study

[5]Unfortunately, Pinotti (2015) does neither state the calculated weights for the $V$ values for the economic predictors nor the RMSPE, the square root of Formula (6) given in (Pinotti, 2015, F216).

[6]The original data used by Pinotti (2015) were downloaded from the website of The Economic Journal.

[7]Weights for the predictors, $V$, for the donor regions, $W$, as well as RMSPE can be found in the first two columns of Tables 1-3.

[8] In all graphics, we follow Pinotti (2015) and shade the time span from 1975 to 1980, when mafia moved to Puglia and Basilicata. Additionally, we also shade the time span used for the SCM estimation.

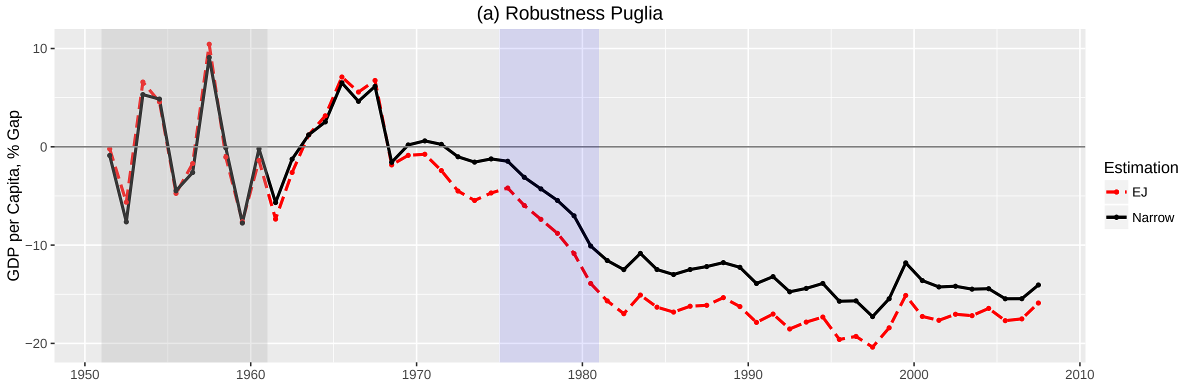


Figure 3: Robustness (a) (Apulia) for replication in a narrow sense

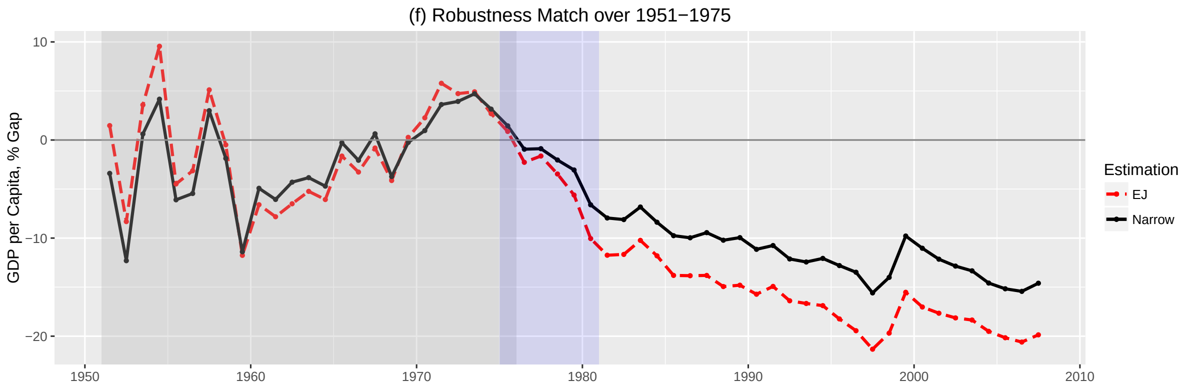


Figure 4: Robustness (f) (Match 1951-75) for replication in a narrow sense

under narrow replication: we find that the gap estimated for Apulia and Basilicata is still quite prominent among the placebo gaps, however, in contrast to the left part of Figure 10 in (Pinotti, 2015, F224), this gap lies no longer completely below all other gaps, raising some doubt whether the effect measured for Apulia and Basilicata is actually significant.

We now turn our attention to the the narrow replication of the robustness studies conducted by Pinotti (2015).[9] For the first robustness study (called '(a) Robustness Puglia') where Puglia is considered separately, Figure 3 shows that, again, the narrow replication produces a slightly smaller gap than the one given by Pinotti (2015). The same holds true for the robustness study called '(f) Robustness Match over 1951-1975' where the time span for fitting the SCM model is prolonged to 1951-75, see Figure 4.

To conclude this section, we follow Pinotti (2015) and consider different measures of economic activity: Figure 5 shows the gaps in electricity consumption per capita, which is a proxy often used for measuring overall economic activity including the shadow economy. Here the results of our narrow replication are only marginally different from those of Pinotti (2015). In particular, our results also imply that the percentage drop in electricity consumption per capita is even larger than the percentage drop in GDP per capita. To explain this substantial drop, Pinotti (2015) considers the shares of value added of agriculture, industry, market services and non-market services. Figure 6 contrasts the

[9]To save space, we concentrate on those results that differ from the ones obtained by Pinotti (2015). Additional results as well as more details, in particular on donor and predictor weights and RMSPEs, are available upon request. Plots for the main variables of interest can be found in Appendix B Figures.

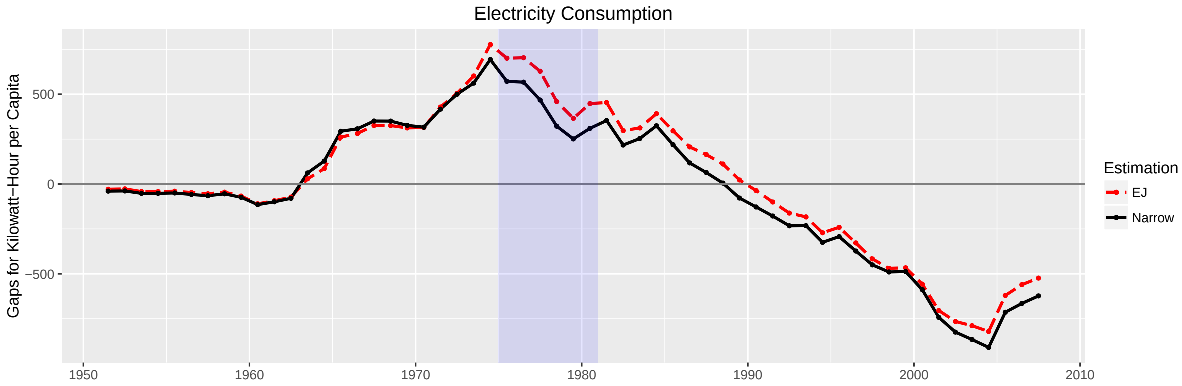


Figure 5: Gaps in electricity consumption for replication in a narrow sense

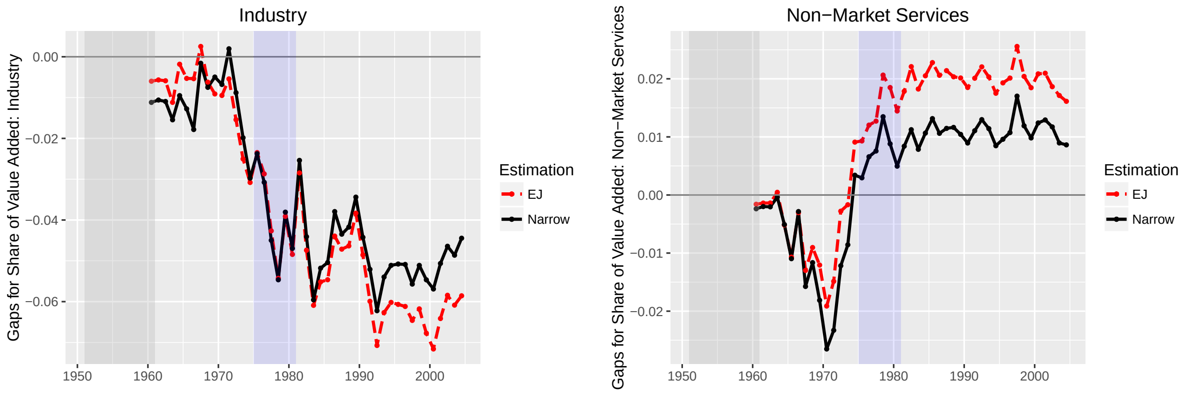


Figure 6: Gaps in share of value added by industry and non-market services for replication in a narrow sense

results of our narrow replication with those of Pinotti (2015): while both agree on a considerable drop of the share of industry coinciding with the mafia outbreak, our narrow replication estimates the increase of the share of non-market services only half as large as Pinotti (2015). Thus, it seems that increased public investment plays a lesser role than presumed by Pinotti (2015).

# 3 Replication Using Generalized Synthetic Control Methods

In this section, we will replicate the studies of Pinotti (2015) in a wider sense by considering several alternative specifications. Our first variant originates from the observation that average GDP per capita attains the largest predictor weight in the original specification as revealed by Table 1. Thus, our first alternative specification, called 'w/o gdp', emerges from removing this economic predictor. As a second alternative, we consider an SCMT approach, i.e. we treat the economic predictors as time series instead of using only their means.[10] Our third alternative specification is motivated by the fact that electric-

[10]Due to all economic predictors but population density starting in 1960 only, this makes only a difference for population density in the main specification. For details and formulas on SCMT as well as MSCM and MSCMT, see Klößner and Pfeifer (2016) and Becker and Klößner (2016a). Note that when

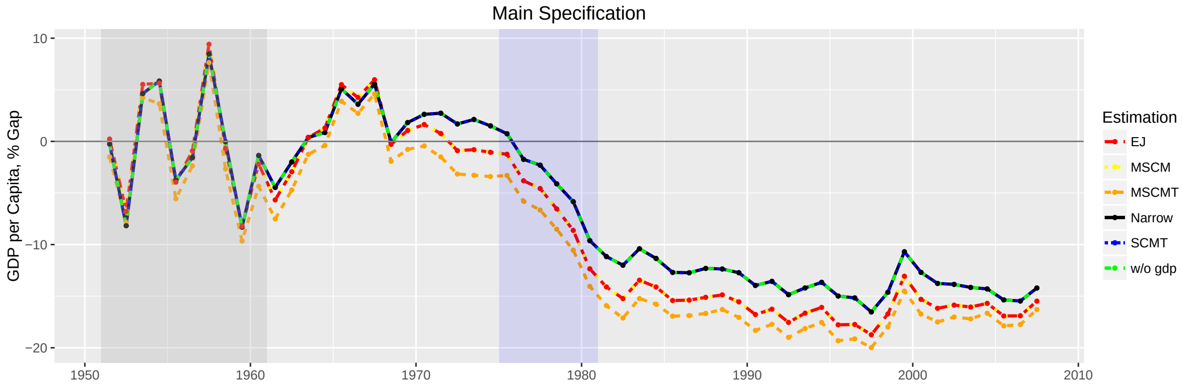


Figure 7: GDP gaps for replication in a wide sense

ity consumption per capita is an important proxy of economic activity and, therefore, is a variable of interest itself. Thus, we add this variable to the dependent variables and employ an MSCM approach. Finally, we also consider an MSCMT approach, where we treat economic predictors as time series and use three dependent variables: GDP per capita, electricity consumption per capita, and the share of value added by industry.[11] Of course, there are many other possibilities for alternative specifications, the ones above were chosen to cover quite a range of structurally different approaches.

Tables 1-3 present predictor and donor weights as well as RMSPE measures for the variables of interest for the time spans 1951-60 and 1951-75. As far as predictor weights are concerned, Table 1 shows that the most important predictor in absence of GDP per capita is the share of value added by industry, except for 'MSCMT' when the share of value added by agriculture and non-market services take that role. The donor weights given in Table 2 coincide for 'Narrow', 'w/o gdp' and 'SCMT',[12] the donor weights for the 'MSCM' replication come very close to those of Pinotti (2015), while 'MSCMT' produces slightly different results, being the only specification for which Emilia Romagna is attributed positive donor weight. With respect to RMSPEs, Table 3 indicates that the differences between the specifications are not very pronounced. Translating the donor weights into gaps for GDP per capita results in Figure 7: here we find for the first time some evidence that the loss in GDP per capita induced by the mafia outbreak might also be estimated marginally larger than by Pinotti (2015), as the gap for 'MSCMT' falls a little bit below the one estimated by Pinotti (2015).

Figure 8 presents the results of the placebo studies for the different variables of interest: the results indicate that the loss in GDP for Apulia and Basilicata is indeed extreme among the placebos, in particular for the specifications 'w/o gdp' and 'MSCMT'. In contrast, the placebo plots for electricity consumption for specifications 'MSCM' and 'MSCMT', together with Figure 10, clearly show that the drop in electricity consumption cannot stem solely from the mafia expanding to Apulia and Basilicata. In line with the findings of Pinotti (2015), Figures 8 and 11 point to a concurrent decrease of the share of value added by industry driving the huge drop in electricity consumption.

treating data as time series, the dependent variable cannot be an economic predictor, as a variable must not be allowed to explain itself.

[11] As explained above, in this case we have to remove GDP per capita and the share of value added by industry from the economic predictors.

[12] These specifications do not always produce identical results, as can be seen from the plots for the robustness studies given in Figures 9-12.

Finally, Figures 9-12 show that all results are fairly robust across the different specifications and robustness studies.

# 4 Conclusion

In this paper, we replicated the study of Pinotti (2015) to estimate the economic costs of organized crime. A narrow replication shows that the results of Pinotti (2015) derived by using Stata are not entirely trustworthy. Our replication results indicate that the loss of economic activity due to the advent of mafia might by slightly smaller than the estimate of 16% given in Pinotti (2015). However, replications in a wider sense provide evidence that the loss in GDP per capita might nevertheless be as large as estimated by Pinotti (2015). Using different more general synthetic control methods, we find both slightly smaller and marginally larger estimates for the drop in GDP per capita, such that overall we can conclude that the estimate of Pinotti (2015) seems to be fairly adequate.

Using the multivariate synthetic control method using time series, we can also shed some light on the question where the loss in GDP per capita may stem from: in line with Pinotti (2015), we find that electricity consumption per capita, often used as a proxy for economic activity including the shadow economy, decreased by even more than 16% in the regions under scrutiny. The results of our new placebo study for electricity consumption allow us to conclude that this drop is not solely caused by the outbreak of mafia activity, as this drop is clearly not significant among the placebos. With respect to the channels through which organized crime impacts the economy, we find that the effect on public investment might be lower than estimated by Pinotti (2015). In line with Pinotti (2015), our estimates point at a significant decrease in industrial investments which causes the loss in economic activity.

# References

**Abadie, Alberto, Alexis Diamond, and Jens Hainmueller**, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

_ **and Javier Gardeazabal**, "The Economic Costs of Conflict: A Case Study of the Basque Country," *The American Economic Review*, 2003, *93* (1), 113–132.

**Ardia, David, Katharine M. Mullen, Brian G. Peterson, and Joshua Ulrich**, *DEoptim: Differential Evolution in R* 2015. version 2.2-3.

**Becker, Martin and Stefan Klößner**, "Computing Generalized Synthetic Controls with the R package MSCMT," July 2016. Working Paper.

_ **and** _ , *MSCMT: Multivariate Synthetic Control Method Using Time Series* 2016. R package version 1.0.0.

**Klößner, Stefan and Gregor Pfeifer**, "Synthesizing Cash for Clunkers: Stabilizing the Car Market, Hurting the Environment," July 2016. Working Paper.

**Pinotti, Paolo**, "The Economic Costs of Organised Crime: Evidence from Southern Italy," *The Economic Journal*, 2015, *125* (586), F203–F232.

**R Core Team**, *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing 2016.

# A Tables

## Main Specification

| | EJ | Narrow | w/o gdp | SCMT | MSCM | MSCMT |
|---|---|---|---|---|---|---|
| Investment Rate | 0.006141563 | 0.010815874 | 0.018351772 | 0.018382899 | 0.006153006 | 0.000000005 |
| Sh. of VA: Industry | 0.464413137 | 0.433223497 | 0.777506641 | 0.778844583 | 0.464330957 | |
| Sh. of VA: Agriculture | 0.006141563 | 0.008839880 | 0.013424505 | 0.013064060 | 0.006153006 | 0.499999948 |
| Sh. of VA: Market Services | 0.013106925 | 0.008839880 | 0.013424505 | 0.013064060 | 0.013206111 | 0.000000088 |
| Sh. of VA: Non-Market Serv. | 0.006141563 | 0.008839880 | 0.013424505 | 0.013064060 | 0.006153006 | 0.499999948 |
| Human Capital | 0.033500548 | 0.087377611 | 0.150443566 | 0.150516278 | 0.033519951 | 0.000000005 |
| Population Density | 0.006141563 | 0.008839880 | 0.013424505 | 0.013064060 | 0.006153006 | 0.000000005 |
| GDP per Capita | 0.464413137 | 0.433223497 | | | 0.464330957 | |

Table 1: Main Specification. Predictor weights.

| | EJ | Narrow | w/o gdp | SCMT | MSCM | MSCMT |
|---|---|---|---|---|---|---|
| Emilia Romagna (EMR) | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 1.5799 |
| Abruzzo (ABR) | 62.4000 | 43.0354 | 43.0354 | 43.0354 | 62.4442 | 67.8025 |
| Molise (MOL) | 37.6000 | 48.9341 | 48.9341 | 48.9341 | 37.5558 | 30.6176 |
| Sardinia (SAR) | 0.0000 | 8.0304 | 8.0304 | 8.0304 | 0.0000 | 0.0000 |

Table 2: Main Specification. Weights of control units in %. Control units with zero weights are omitted.

| | EJ | Narrow | w/o gdp | SCMT | MSCM | MSCMT |
|---|---|---|---|---|---|---|
| GDP per Capita 1951–60 | 130.141 | 128.949 | 128.949 | 128.949 | 130.141 | 139.868 |
| GDP per Capita 1951–75 | 121.189 | 122.873 | 122.873 | 122.873 | 121.206 | 144.928 |
| kWh per Capita 1951–60 | 55.375 | 63.413 | 63.413 | 63.413 | 55.375 | 59.915 |
| kWh per Capita 1951–75 | 316.406 | 300.348 | 300.348 | 300.348 | 316.262 | 286.449 |
| Sh. of VA: Industry 1951–60 | 0.006 | 0.011 | 0.011 | 0.011 | 0.006 | 0.007 |
| Sh. of VA: Industry 1951–75 | 0.013 | 0.014 | 0.014 | 0.014 | 0.013 | 0.015 |

Table 3: Main Specification. RMSPE of GDP & electricity consumption per capita and share of value added by industry for periods 1951–60 and 1951–75.

# B Figures

## Placebo Plots

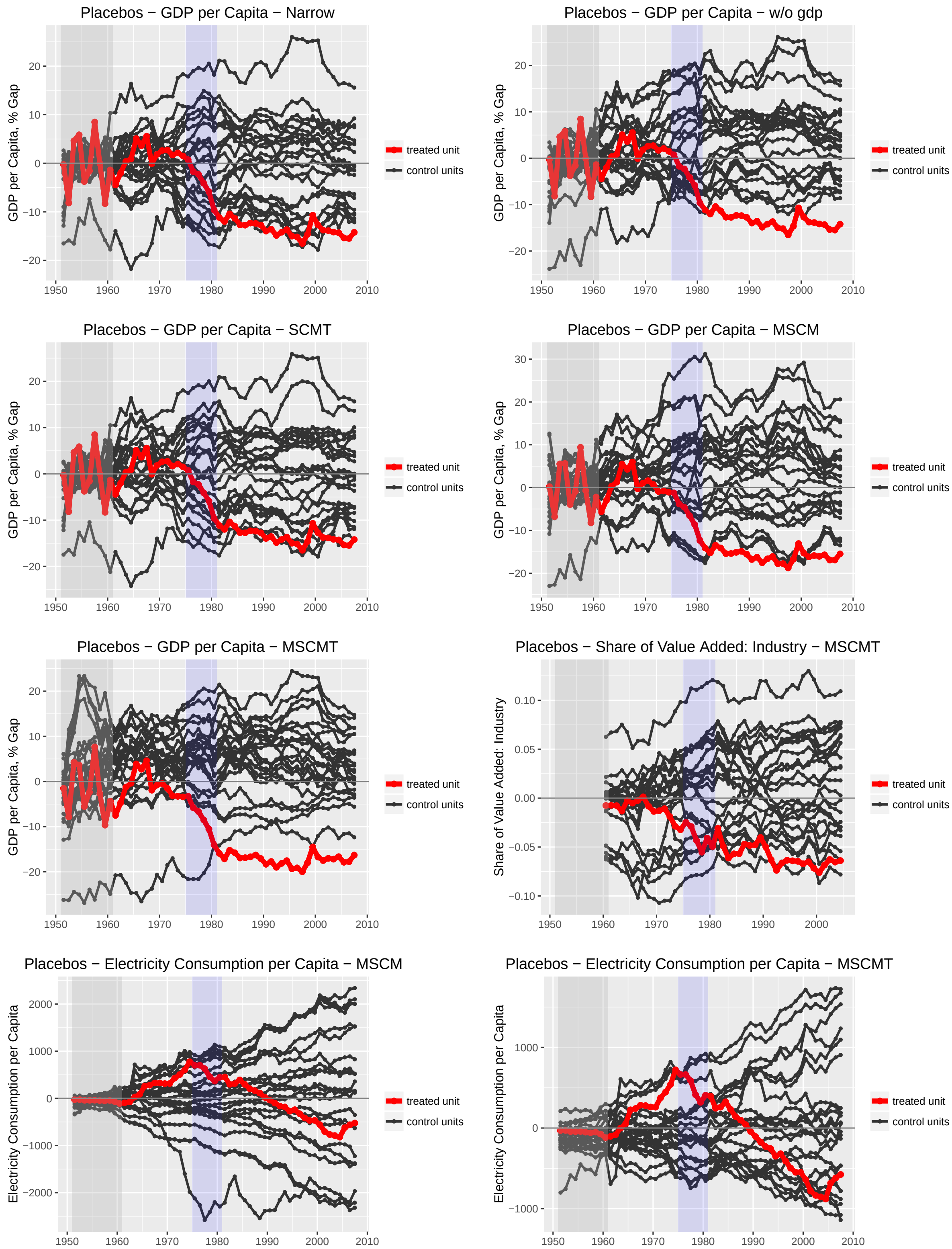


Figure 8: Placebo plots for different dependent quantities and specifications

# GDP per Capita, % Gap

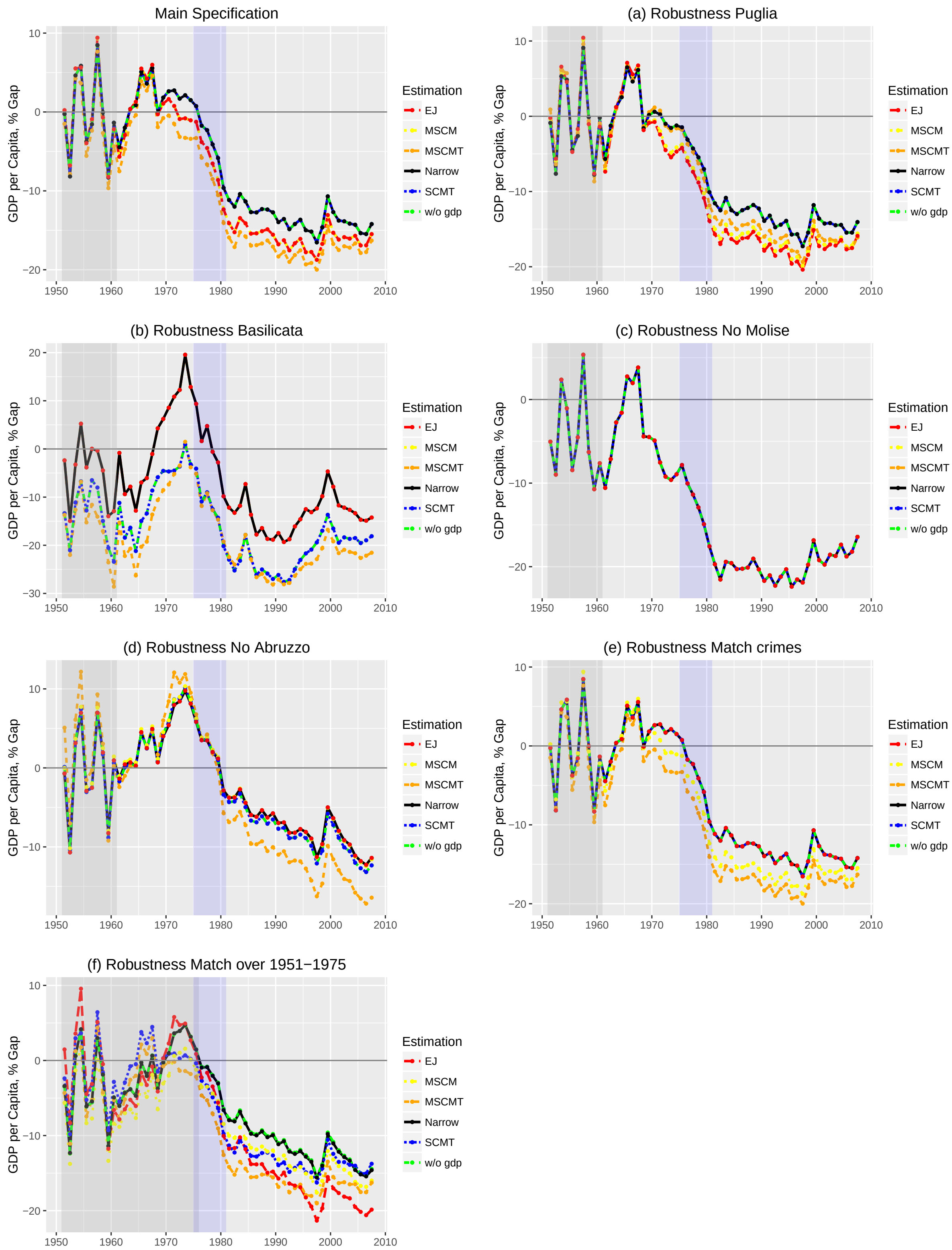


Figure 9: Gaps for GDP per Capita, % Gap

# Electricity Consumption per Capita

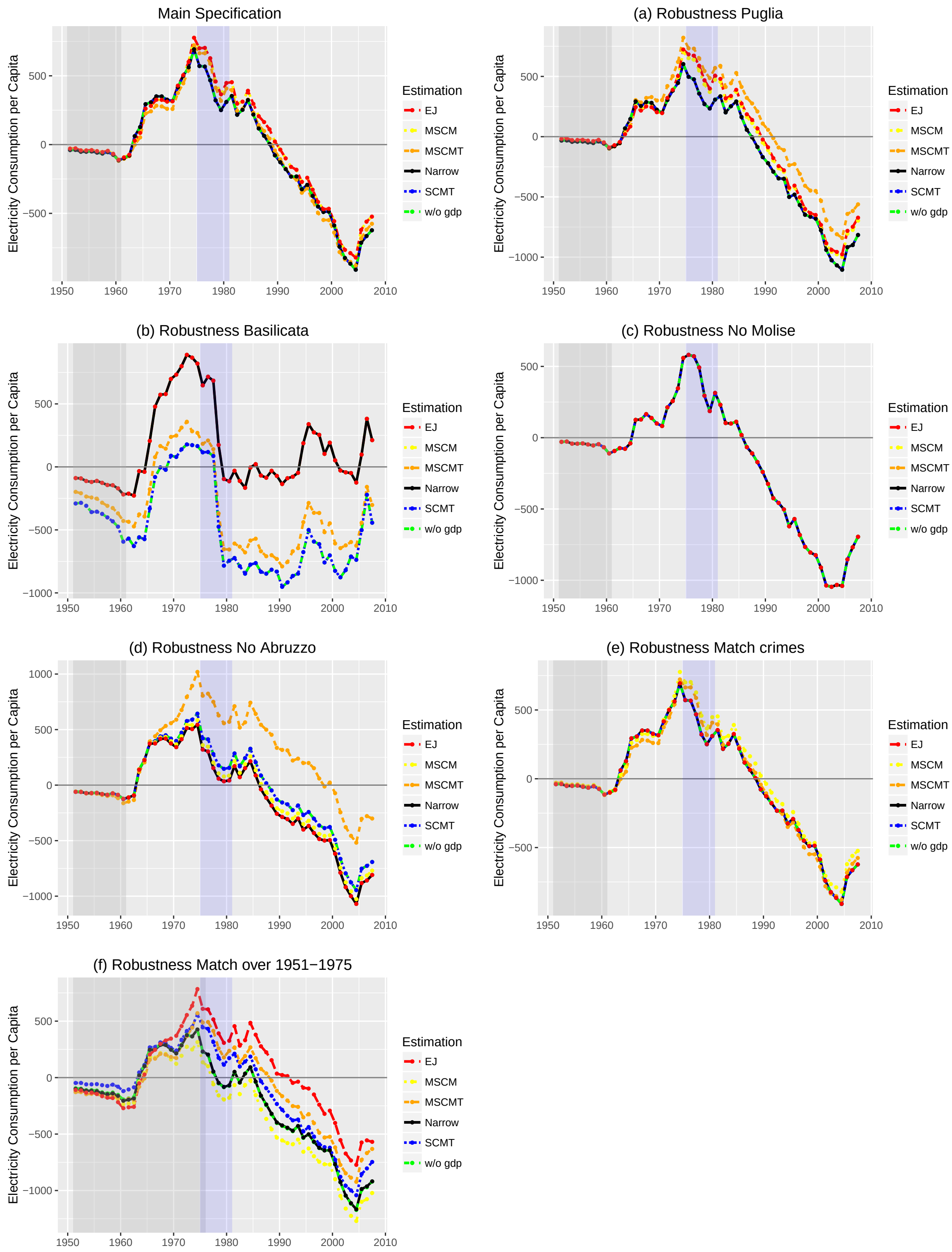


Figure 10: Gaps for Electricity Consumption per Capita

# Share of Value Added: Industry

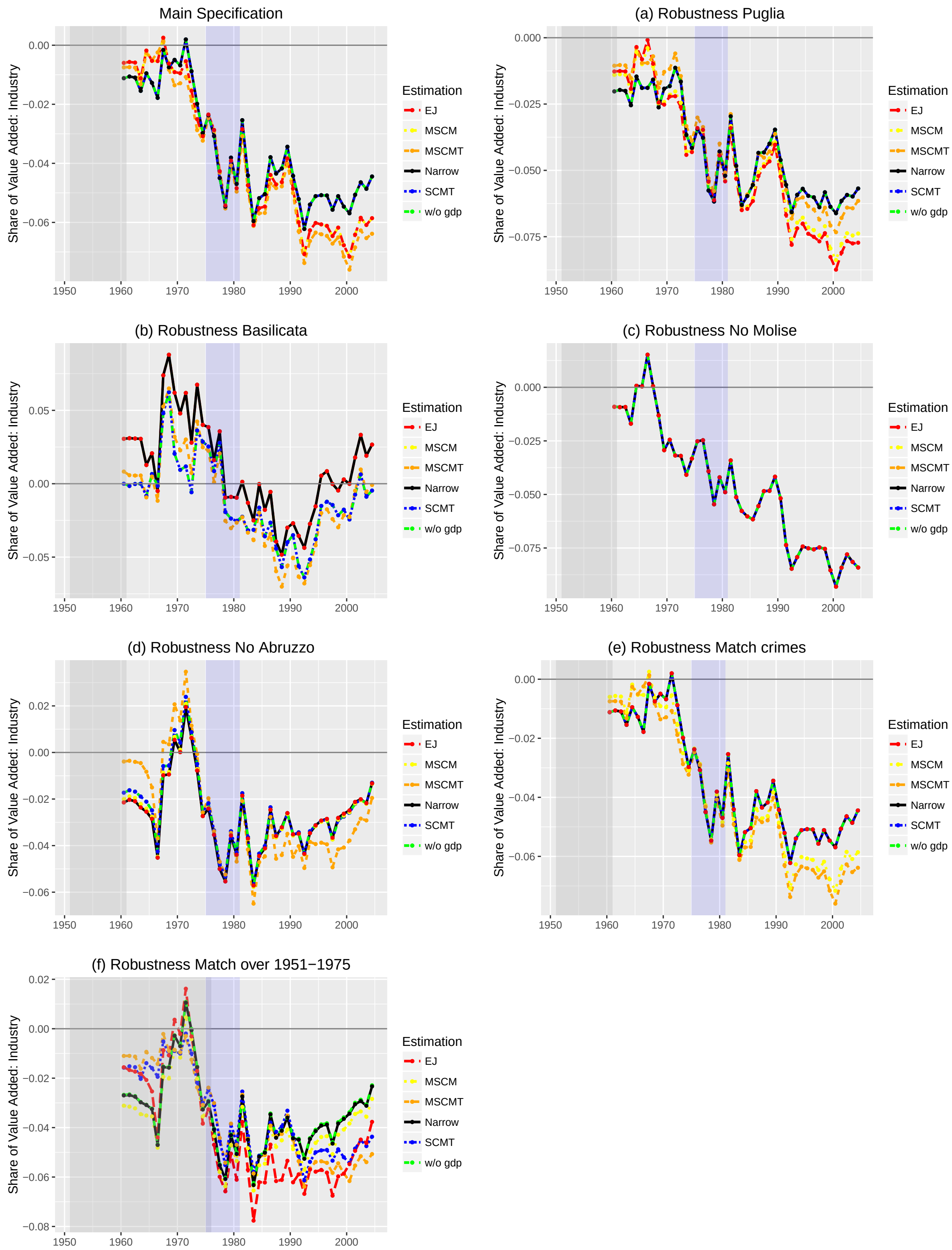


Figure 11: Gaps for Share of Value Added: Industry

## Share of Value Added: Non-Market Services

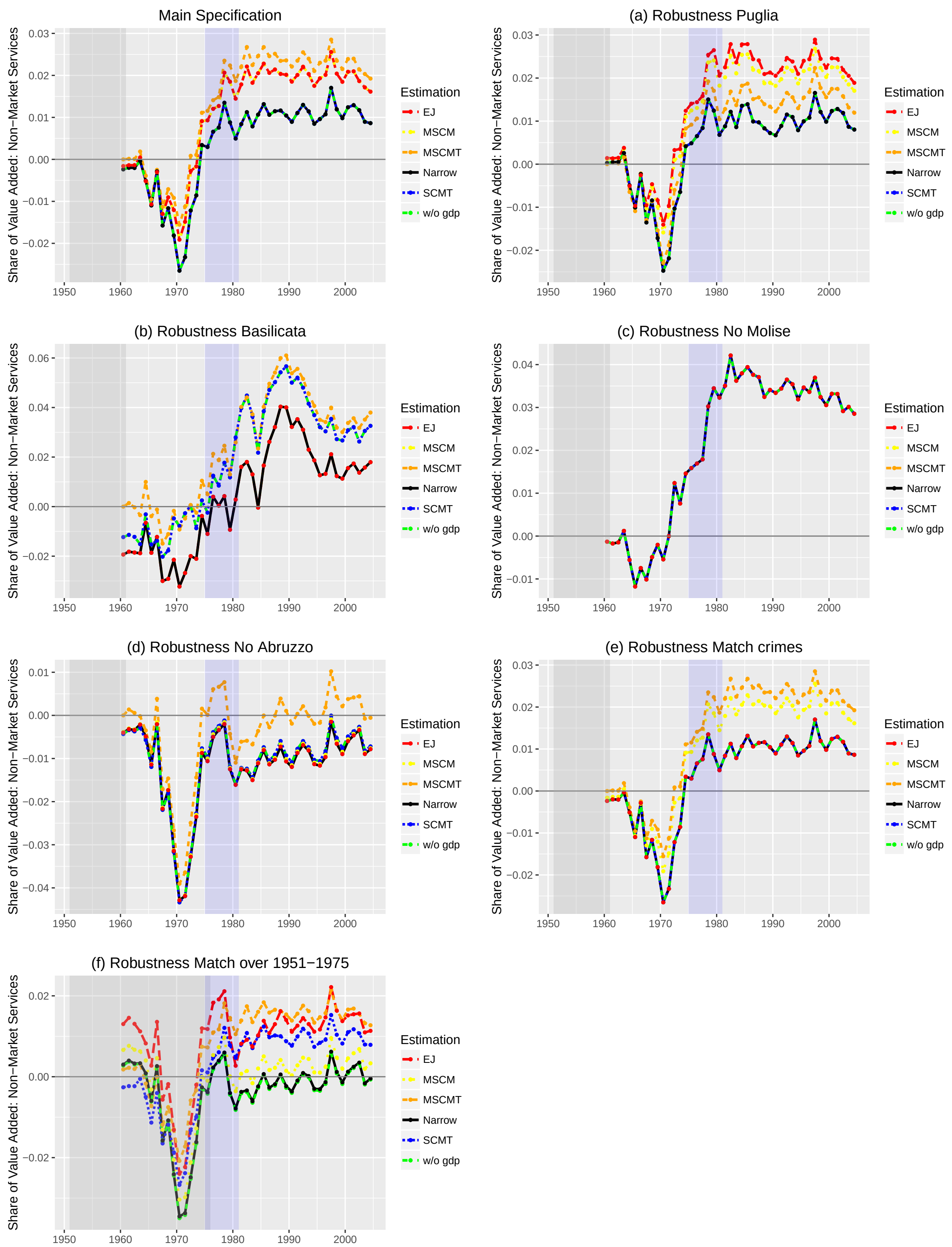


Figure 12: Gaps for Share of Value Added: Non-Market Services