

Income and Productivity Trends of Guangdong: a Data-driven Case Study of the Greater Bay Area*

Jingwei Li[†]

Lucas Shen[†]

Xuyao Zhang[†]

October 2022

Abstract

In 2017, the Chinese government announced plans for the development of the Greater Bay Area, a world-class city cluster in southern China. Our study statistically compares the income and productivity of Guangdong, home to nine bay area cities, to other mainland provinces since the announcement. We observed a decline in both measures in Guangdong due to the influx of people since the announcement. We attribute this increase in population to 1) higher employment in the low productivity sectors like construction, wholesale, and retail sectors and 2) key policies aimed at attracting talents from other mainland provinces.

Keywords: Greater Bay Area, Demographic Trends, Macroeconomic Effects, Regional Economic Activity, Synthetic Control Method

JEL Classification: F15, J11, O53, R11

*We are grateful to our colleagues and Michael Makovi for helpful comments. We also thank Zhou Jingwei for additional research assistance. Maps and other charts are neutral with regard to national sovereignty and jurisdictional claims. Other usual disclaimers apply.

[†]Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore. JL: jingweil@nus.edu.sg LS: lsys@nus.edu.sg. XZ: sppzhhx@nus.edu.sg.

1 Introduction

The Guangdong-Hong Kong-Macao Greater Bay Area (GBA), consisting of nine Guangdong cities (also known as Pearl River Delta region) and two Special Administrative Regions, is the new bay area and a megalopolis in the making in China. The proposal of the GBA, signed into agreement in 2017, comprises a multitude of policies that aims to provide the region with conditions suitable for high-quality economic growth.

Our study aims to examine the impact of the 2017 announcement and signing of the 2017 GBA development framework on Guangdong's economy. Excluding the Hong Kong and Macau Special Administrative Regions, all the nine remaining cities in the GBA reside in the Guangdong mainland province and constitute most of Guangdong's economy.¹ We evaluate basic economic indicators before and after the 2017 GBA, focusing on Guangdong province as the epicentre of the GBA, to assess the impact of the 2017 GBA announcement.

The Greater Bay Area The GBA consists of the Hong Kong Special Administrative Region, the Macao Special Administrative Region, and nine Guangdong cities in the Pearl River Delta (PRD) region (Guangzhou, Shenzhen, Zhuhai, Foshan, Huizhou, Dongguan, Zhongshan, Jiangmen, and Zhaoqing), covering a total urban area of 56,000 square kilometres with a combined population of approximately 70 million.

The GBA region's development was first introduced in 2015 in *The Vision and Actions for Jointly Building the Silk Road Economic Belt and the 21st Century Maritime Silk Road* jointly issued by the National Development and Reform Commission, the Ministry of Foreign Affairs, and the Ministry of Commerce. In 2017, a proposal to “draw up a plan for developing a city cluster in the Guangdong-Hong Kong-Macao Greater Bay Area” was written in a government report. Since then, the development of the GBA has expanded to one of national focus. In 2019, China released the “Outline Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area.” The outline stipulates a short-term plan for 2022 and a long-term project for 2035, establishing goals

¹According to Guangdong Statistical Yearbook, in 2020, the Pearl River Delta (PRD) region (with municipalities including Guangzhou, Shenzhen, Zhuhai, Foshan, Huizhou, Dongguan, Zhongshan, Jiangmen, and Zhaoqing) contributed 80.5% of the GDP in Guangdong province and includes 62.0% of the provincial population.

for the two phases.

This study uses the year 2017 as the year of intervention, as it marks the official launch of the GBA after the signing of the Framework Agreement on Deepening Guangdong-Hong Kong-Macao Co-operation in the Development of the Bay Area on 1st July 2017. [Figure 1](#) shows a map with China, Guangdong as our main province of interest, and the nine PRD cities within Guangdong.²

Empirical approach To examine the impact of the 2017 GBA announcement, our case study focuses on the Guangdong province as the epicentre of the GBA. We use the synthetic control method (Abadie and Gardeazabal 2003; Abadie et al. 2010) as a quantitative approach to a case study where only one unit (Guangdong province) experienced an intervention. Implementing the synthetic control method allows us to impute the counterfactual trend of Guangdong's Per Capita GDP and Population density using the weighted combination of the other mainland provinces. To the extent that the counterfactual analysis is valid, the difference between Guangdong's trend and the counterfactual trend reveals the impact of the 2017 announcement.

Our findings Our first set of findings is for Per Capita GDP. We find robust and statistically significant quantitative evidence that after the 2017 GBA announcement, Guangdong's Per Capita GDP falls relative to the counterfactual trend. A placebo exercise reveals that this drop in Guangdong's Per Capita GDP is absent in the other mainland provinces, suggesting that we are not simply capturing an economic shock common to the other China regions. We further perform a set of tests to show that the fall in Guangdong's Per Capita GDP, relative to the other counterfactual trend, is not sensitive to the subset of mainland provinces used to construct the counterfactual, not sensitive to restriction the synthetic control estimation to a shorter period, and not sensitive to the set of macroeconomic indicators used to predict the counterfactual trend.

Our second set of findings reveals a rise in population as the reason for the observed fall in synthetic control estimation suggests that Guangdong experienced a sharp increase in population density following the 2017 GBA announcement. This is due to an increase in Guang-

²The Framework was signed by the Chief Executive of the Hong Kong Special Administrative Region, the Chief Executive of the Macau Special Administrative Region, the Chairman of the National Development and Reform Commission, and the Governor of Guangdong Province, witnessed by President Xi Jinping. The [Supplementary Materials](#) provides additional notes on Guangdong and the GBA.

dong's population is similarly robust and not present in the other mainland provinces. Since the standard productivity measure of Per Capita GDP is GDP divided by population size, our results suggest the fall in Guangdong's productivity is attributable to an increase in population.

In the discussion and further analyses, we focus on labour productivity instead, where we find a similar decline in Guangdong's productivity and suggest that the observed decline in economic performance is not unique to the Per Capita GDP measure. We also link our findings on the influx of population to GBA policies involving talent recruitment at the provincial level, including benefits such as housing subsidies. Overall, our results suggest a broad-stroke evaluation of an uprising bay area is not always straightforward because the inception of the bay area itself leads to compositional changes in the economy.

Discussion To refrain from statistical exercises bordering on multiple hypothesis testing, we have restricted our analyses to the core and fundamental indicators of economic performance. Our focus on population (density) stems from the fact that the decline in both Per Capita GDP and labour productivity leads to a potentially misleading conclusion that the GBA had a negative effect. The rise in population of Guangdong around the same time provides evidence for this. What we are capturing stems from demographic changes rather than any fundamental decline in output. The reason we do not study the impact of GBA on just GDP is because such measures are not normalized and less comparable across provinces (Abadie 2020). When we study demographic changes in population, these are normalized by land area so that we are looking at the normalized measure of population density instead of absolute levels of population.

We emphasize that while our sample period includes the year 2020, this should not be a primary concern in how we draw our conclusion. First, the Covid-19 pandemic is a worldwide shock and does not impact only Guangdong (or any other province). Second, the key demographic change in population (density) that explains the decline in Guangdong's Per Capita GDP and labour productivity occurs prior to 2019–2020. In any case, for all our key results, we put them through a battery of sensitivity tests. These help rule out how timing of the intervention or the makeup of the synthetic Guangdong as a counterfactual to Guangdong, which potentially captures differential

effects of the pandemic, can explain our results.

Urbanization and policies that aim at promoting infrastructure for regional connectivity, as well as the attraction of capital and labour talent, should boost economic growth. The GBA is one such plan, and is comparable with other bay areas worldwide (including the New York, Tokyo, and San Francisco bay areas). Based on an early evaluation of fundamental indicators of economic performance, we do not observe such an economic growth spurt. Instead, we find a decline in Per Capita GDP and labour productivity. We draw on fundamental economic reasoning and show that any interpretation that the GBA was detrimental to economic growth is spurious. What we are capturing is a rapid and dramatic shift in population flows into the cities of Guangdong that led to an estimated decline in the measures normalized by population. There, in the short run, certain fundamental indices may become less reliable due to such a dramatic demographic shift. Such shifts are not unique to the GBA and apply to all regions with such urbanization and regionalization plans. Had we not exposed the underlying demographic changes, we would have spuriously attributed the GBA as detrimental to the economy of Guangdong, which is not the case. Our analyses occur at the provincial level, taking Guangdong as the epicentre of the GBA. This does not preclude heterogeneous effects in terms of early benefits. In particular, [Zhang et al. \(2022\)](#) finds that the city of Foshan is an early beneficiary of the GBA with increases in retail sector job creation from the increase in population. In ongoing work, we aim to study more closely the patterns of talent migration from both within and from outside of the mainland provinces.

Related studies Since the Eleventh Five Year Plan (2006-2010), urban clusters have become the main driver of China’s economic growth. The most successful examples include the Yangtze River Delta and the Pearl River Delta. Researchers have then dived into this area to examine various economic factors in the clusters, such as industrial development ([Lee and Lin 2020](#); [Tang 2021](#); [Liu and Zhang 2021](#)) innovation ([Sheng et al. 2019](#); [Hu and Kim 2022](#); [Chow et al. 2020](#)), infrastructure ([Weng et al. 2020](#); [Hui et al. 2020](#); [Ding et al. 2019](#)) and FDI ([Wen 2014](#); [Zhang et al. 2022](#)).

On the one hand, studies have revealed that urban agglomeration is an important factor for innovation ([Sheng et al. 2019](#)) and the development of financial industries ([Liu and Zhang 2021](#)).

Ding et al. (2019), Hui et al. (2020), and Wen (2014) have discussed the importance of infrastructure and FDI to boost economic growth in the clusters. On the other hand, Xu and Jiao (2021) argued that the expansion of the urban population weakens the positive effect of urbanization. This finding is similar to ours, as we have shown that the rising population in Guangdong province has lowered the region's productivity.

Of the aforementioned studies, only a subset focuses on the GBA (Hu and Kim 2022; Liu and Zhang 2021; You et al. 2022). From other research field, Ni et al. (2020), Loo and Bo (2018), Bin and Lu (2020) have focused on regional integration in GBA. They have examined the degree of integration from spatial patterns, rail transit networks, and the establishment of free trade pilot zones. Zhang et al. (2022) contributes to this area from the angle of FDI. Since there is a differential development plan for Core and Node cities in GBA, the study examines whether manufacturing-related FDI is drawn to the Node cities and services-related FDI is drawn to the Core cities.

Our study fills the literature gap by studying the impact of the 2017 announcement on economic growth and its relations with a growing population, which may provide better insights into the area's talent attraction policies and innovation efficiency.

Section 2 describes empirical methodology we use. Section 3 describes our data. Section 4 and Section 5 presents the results for Per Capita GDP and Population density. Section 6 extends our analyses to productivity and discusses talent attraction policies. Section 7 concludes.

2 Methodology

The common methodological approach when studying the implementation or announcement of a policy is to use the difference-in-differences design. In such an approach, the counterfactual to the provinces with a policy intervention is the weighted outcome of the other provinces who would otherwise share the same trend had the policy intervention not been implemented.³

³See Abadie (2020) for a more technical discussion of different weights arising from difference-in-differences vs. synthetic control estimation.



a China, Guangdong, and the PRD cities

b Guangdong

Figure 1. Left panel shows the nine Pearl River Delta cities in the Greater Bay Area (shaded) as part of China. The broader outline indicates the Guangdong province. Right panel shows Guangdong and the nine Pearl River Delta cities (Guangzhou, Zhuhai, Shenzhen, Foshan, Huizhou, Dongguan, Jiangmen, Zhongshan, and Zhaoqin) plus the Hong Kong and Macau Special Administrative Regions as part of the Greater Bay Area.

[[Figure 1](#) near here]

Our Greater Bay Area study exists in a context that differs from the standard difference-in-differences setup in at least two ways. First, we only have one province, Guangdong, under the Greater Bay Area instead of many such provinces from which we can average over. Second, and more importantly, the standard parallel trends assumption needed to recover the average effect of the Greater Bay Area is unlikely to be satisfied given that Guangdong is already one of the larger economies in China and was on a different trajectory than the average mainland provinces even before the 2017 announcement (e.g., [Lee and Lin 2020](#); [Sheng et al. 2019](#); [Weng et al. 2020](#) and [Supplementary Materials](#)). Hence, the empirical approach in our study is the synthetic control method which does not require the parallel trend assumption and is designed for case studies with only a single “treated” unit, as in our case.

Below, we describe in more details our implementation of the synthetic control method ([Abadie and Gardeazabal 2003](#); [Abadie et al. 2010](#); [Abadie 2020](#)), where our aim is to impute the counterfactual trend of Guangdong using a weighted combination of the other mainland provinces.

The objective of our synthetic control method implementation is to create a synthetic Guangdong Province using the weighted average of the other Mainland China provinces to resemble the

characteristics of Guangdong Province. The characteristics of our interest are in two areas: macroeconomic performance and human capital. [Table A2](#) in the [Supplementary Materials](#) section lists the macroeconomic variables we use in our synthetic control estimation. Let Y and X represent the outcome variable and predictors, respectively. The main outcome variable in our model is Per Capita GDP, and the rest of the variables are predictors.

Let P be the number of provinces in the control group (there are 30 Mainland China provinces other than Guangdong). With a slight abuse of notation, let P also be the set of provinces. Let w_p be the weight assigned to each province, $p \in P$, with $0 \leq w_p \leq 1$ and $\sum_{p \in P} w_p = 1$.⁴

Let X_G and X_P be the set of predictors for Guangdong Province and the control provinces, respectively. We want to minimize the sum of the differences between the pre-policy intervention characteristics of Guangdong (X_G) and synthetic Guangdong ($X_P W$) by choosing the set of weights W :

$$\underset{W}{\text{Min}} (X_G - X_P W)' V (X_G - X_P W), \quad (1)$$

where V are the weights for each of the predictor variables, and V^* is calculated to minimize the difference in the outcome variable between Guangdong and synthetic Guangdong in the pre-intervention period:

$$\underset{V}{\text{Min}} \sum_{t \in T_0} (Y_{Gt} - \hat{Y}_{Gt})^2, \quad (2)$$

where $\hat{Y}_{Gt} = \sum_{p=1}^P w_p^* Y_{pt}$ and T_0 is the time period before the policy shock.

The sets of weights W and V then together constitute the synthetic control of Guangdong, and the impact of the policy is calculated as the difference between Guangdong and synthetic Guangdong:

$$\tau_{Gt} = Y_{Gt} - \hat{Y}_{Gt}, \quad (3)$$

for each $t > T_0$.

⁴We drop Hong Kong and Macau from the pool of donors because they do not operate in a similar system from the mainland provinces, and more importantly, receive the same shock as Guangdong as part of the GBA. See the [Supplementary Materials](#) for more details.

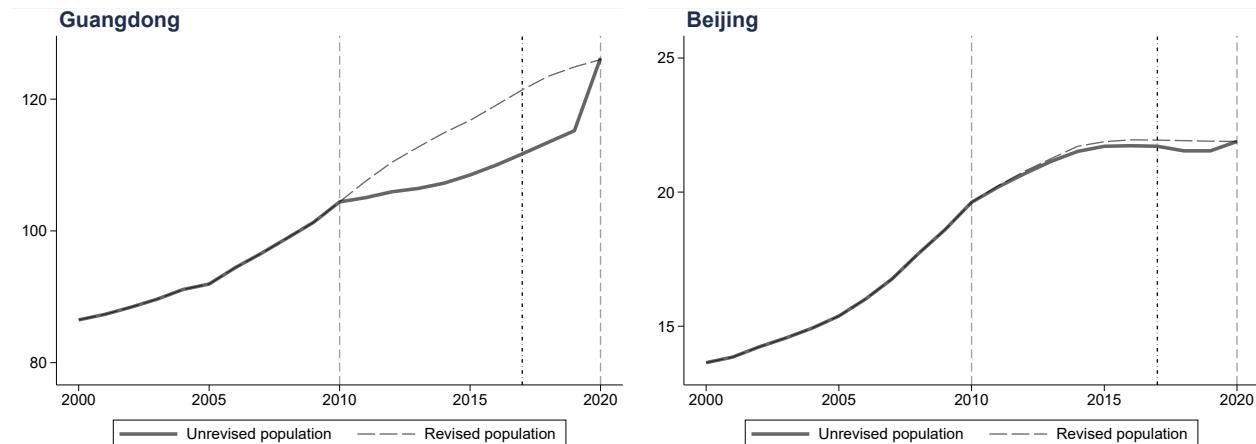


Figure 2. Population path plots for Guangdong and Beijing. Unrevised and revised population plotted are the raw data from the panel for 2000–2020. Dashed vertical lines are for the years 2010 and 2020, which are the endpoints for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

[Figure 2 near here]

3 Data

For this study, we collect data on macroeconomic indicators from various public sources, including the National Bureau of Statistics of China, China Statistical Yearbooks, provincial yearbooks, and China Labour Statistical Yearbooks. The sample period is over 21 years, from 2000–2020.

In this section, we highlight the population data used in this study. As listed in Table A1 of the [Supplementary Materials](#), the data for 2000, 2010, and 2020 are from the population census conducted every ten years. The sampling method for the data collection is the main difference between the census years and the non-census years. Census takers collected census data through a door-to-door method from all the households. To verify the accuracy of the census, for example, after the census in 2020, 32,000 households were randomly selected from 141 counties among the 31 Mainland China provinces. 0.05% of the households were not visited by the census takers, lower than the 3% international standard.⁵

For the non-census years, a random sampling of households among the 31 provinces is conducted. The coverage is around 0.1% of the total population. At the mid-point between two census

⁵http://www.stats.gov.cn/tjsj/zxfb/202105/t20210511_1817274.html accessed on 15 Dec 2021.

years, i.e., 2005 and 2015, the coverage is increased to 1% of the total population.

According to the National Bureau of Statistics, population data from 2011 to 2019 are revised based on the 2020 census results. We can see a clear data smoothing in [Figure 2](#).⁶ Population plots for the remaining provinces are in the [Supplementary Materials \(Figure A2–Figure A7\)](#). The unrevised population data is collected from the China Statistical Yearbooks or provincial statistical yearbooks, while the revised data is collected from the National Bureau of Statistics. We use the unrevised data in this study from 2011 to 2019.⁷ Due to the unavailability of yearbooks for some provinces and to ensure consistency, for 2001 to 2009 data, we have collected the data from the National Bureau of Statistics in 2021. Thus, for this period, the data is revised.

We acknowledge concerns regarding our choice in using the unrevised population in non-decennial census years. We emphasize that all population, both unrevised and revised, originates from the official statistics rather than through our own estimates. The only difference between the unrevised and revised population numbers, according to our observations ([Figure 2](#) and [Figure A2–Figure A7](#) in the [Supplementary Materials](#)) is that the revised population smooths out kinks in the unrevised data. [Figure 2](#) demonstrates this for both Guangdong and Beijing. Changes that are smoothed out in this manner are less likely to show up in estimations, including our synthetic control estimation. This is the primary reason we perform our analyses with the unrevised data. [Figure A2–Figure A7](#) in the [Supplementary Materials](#) show the smoothing out of changes for all mainland provinces.

4 Potential Effect on Per Capita GDP

[Figure 3](#) shows the general trend of Per Capita GDP using the raw data before and after the GBA announcement in 2017. To estimate whether the dip in Guangdong's Per Capita GDP can be attributed to the announcement, we use the synthetic control method described in [Section 2](#).

⁶http://www.stats.gov.cn/tjsj/zxfb/202105/t20210512_1817360.html accessed on 15 Dec 2021.

⁷For example, data for 2011 are collected from 2012 statistical yearbooks. Since the data could not be revised in the yearbooks, it will be considered the unrevised data, which is collected by the random sampling method.

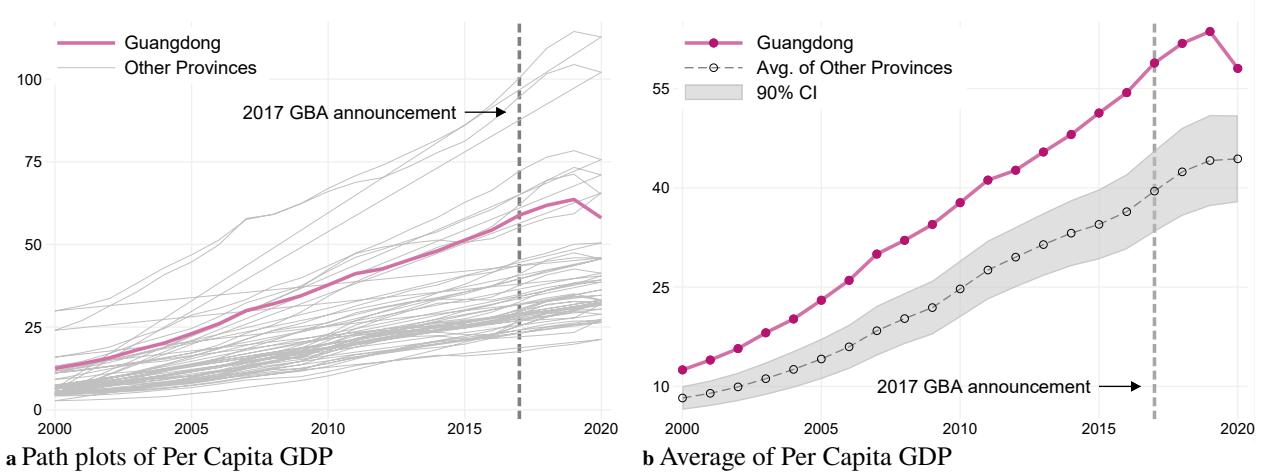


Figure 3. Panel a shows the raw path plots of the province Per Capita GDP. Gray lines are mainland provinces other than Guangdong. Panel b shows the path plot of Guangdong and the average of the other provinces. Shaded gray area is the 90% confidence interval. Dashed vertical line indicates the 2017 announcement of the GBA.

The synthetic control method imputes the counterfactual trend of Guangdong, using the weighted combination of the other mainland provinces before the GBA announcement, to construct synthetic Guangdong. To the extent that synthetic Guangdong is a good approximation, the difference between Guangdong and synthetic Guangdong reveals the impact of GBA on the Guangdong economy.

[Figure 3 near here]

In addition to exposing trends, Figure 3 also demonstrates that, while Guangdong is an outlier in terms of Per Capita GDP (panel b), a convex combination of other mainland provinces as a synthetic counterfactual for Guangdong can still be constructed since its trajectory still falls between the trajectory of many other provinces (panel a) (Abadie 2020).

In this section, we report quantitative evidence that the drop in Guangdong’s Per Capita GDP coincides with the GBA announcement in a way that is not reflected in the other mainland provinces.

4.1 Results

Figure 4 show the lists of donor provinces and predictor variables we use in the synthetic control estimation. Each set of donor province weights corresponds to a potential synthetic control unit

for Guangdong. The left panel of [Figure 4](#) shows the set of weights we use, which minimizes the pre-GBA characteristics of Guangdong and the pre-GBA characteristics of the counterfactual synthetic Guangdong ([Equation \(1\)](#)). As evident from the left panel of [Figure 4](#), only four mainland provinces contribute to synthetic Guangdong.⁸

[[Figure 4](#) near here]

For the pre-GBA characteristics in the synthetic control estimation, the seven economic indicators we use are (i) the primary industry value (scaled by GDP), (ii) the secondary industry value (scaled by GDP), (iii) the tertiary industry value (scaled by GDP), (iv) annual (urban) household income per capita, (v) the number of international visitor arrivals (scaled by GDP), (vi) the (urban) unemployment rate, and (vii) the average annual wage of staff and workers (see [Table A2](#) of the [Supplementary Materials](#)). The right panel of [Figure 4](#) then shows the variable weights for the seven predictor variables where the value of the tertiary industry has the highest weight for predicting Per Capita GDP, followed by wage and the unemployment rate.

[Table 1](#) shows how closely the pre-GBA characteristics of Guangdong matches the pre-GBA characteristics of pre-GBA synthetic Guangdong. We also report the pre-GBA match in characteristics between Guangdong and three special cases of synthetic Guangdong as a benchmark. The first special case is where each of the 30 mainland provinces receives equal weight in constructing synthetic Guangdong. The second special case is when the nearest neighbour, the province with the highest optimized weight, gets the full weight while all other provinces receive zero weight. The third special case is for the second nearest neighbour—the province with the second-highest optimized weight. Synthetic Guangdong constructed using the synthetic control method has the

⁸The sparsity of synthetic weights is a feature of the synthetic control method ([Abadie 2020](#)). Guangdong, Beijing, Shandong, and Jiangsu are in the Eastern Coastal Region, which is the most developed area in China. In terms of GDP, Guangdong, Jiangsu, Shandong, Zhejiang, and Henan are the top five provinces in 2021. Thus, it is not surprising that these four provinces form synthetic Guangdong. One concern is that any single non-zero donor province (e.g., Henan, Beijing, Shandong, and Jiangsu) or other donor provinces in the initial donor pool artificially affects our results. For instance, Henan experienced heavy flooding in 2021. While this is not in our sample period, such events create an artificial depression of Per Capita GDP in the synthetic unit and lead to an artificially low decline in the estimated impact of the GBA. Other events that certain provinces encounter may lead to artificially high estimated increases. To deal with this, we perform leave-one-out sensitivity analyses that show that omitting any single donor province from the synthetic control estimation does not substantially affect our results ([Section 4.2](#)).

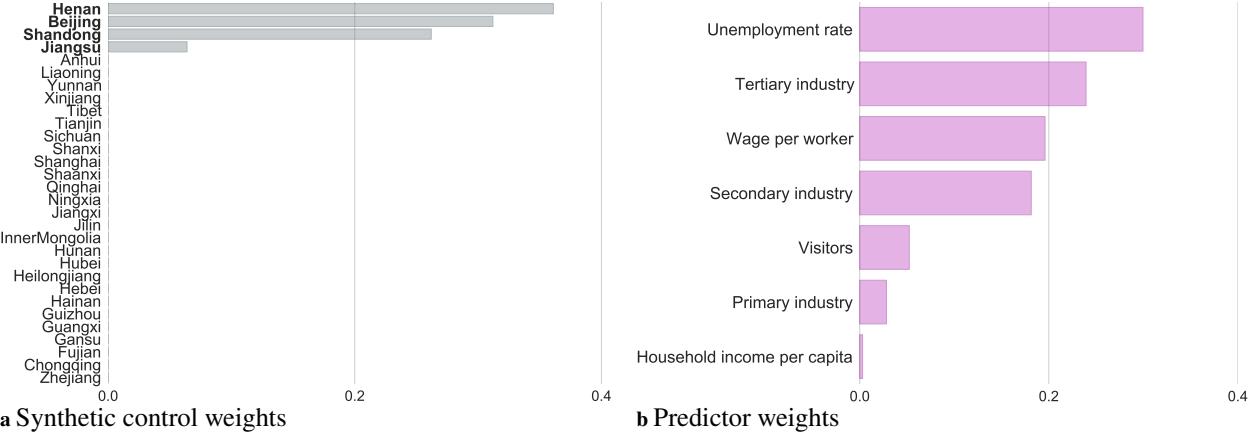


Figure 4. Weights to construct Per Capita GDP of synthetic Guangdong. Left panel are the synthetic control weights, where weights are W^* from Equation (1) chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Right panel are the predictor variable weights V^* from Equation (2).

Table 1. Balance of predictors of Per Capita GDP before the 2017 GBA announcement.

	Weight	Guangdong	Synthetic Guangdong	Average	Henan (nearest neighbour)	Beijing (next nearest neighbour)
Primary industry	0.028	0.059	0.096	0.132	0.155	0.011
Secondary industry	0.181	0.479	0.453	0.477	0.521	0.239
Tertiary industry	0.239	0.462	0.472	0.447	0.336	0.699
Household income per capita	0.003	19,214	16,935	14,217	11,714	26,354
Visitors	0.052	0.018	0.016	0.015	0.004	0.036
Unemployment rate (urban)	0.3	2.601	2.701	3.591	3.159	1.447
Annual wage	0.196	30,861	30,277	26,547	20,079	51,925

Each row is a predictor used to match the Per Capita GDP of the 30 mainland provinces to Guangdong. Column 1 is the variable weight. Column 2 is the mean values for Guangdong. Column 3 is the mean predictor values for synthetic Guangdong ($X_P W^*$ from Equation (1)). Columns 4, 5, and 6 are special cases of the synthetic control. Column 4 is where all 30 mainland provinces get equal weights (W is a vector with all elements $1/30$). Column 5 is when the nearest neighbour gets full weight (W all zeroes except with weight 1 for Henan, see Figure 4). Column 6 is when the second nearest neighbour gets full weight (W all zeroes except with weight 1 for Beijing).

best overall match.

[Table 1 near here]

The synthetic control results in Figure 5 confirm that Guangdong's Per Capita GDP drops after the 2017 GBA announcement. The left panel shows the path plot of Per Capita GDP for Guangdong and for synthetic Guangdong. Both paths mostly coincide before the announcement and then diverge after, with Per Capita GDP lower for Guangdong than for synthetic Guangdong. The right panel shows the difference over time, with differences clustering around zero before diverging after the announcement.

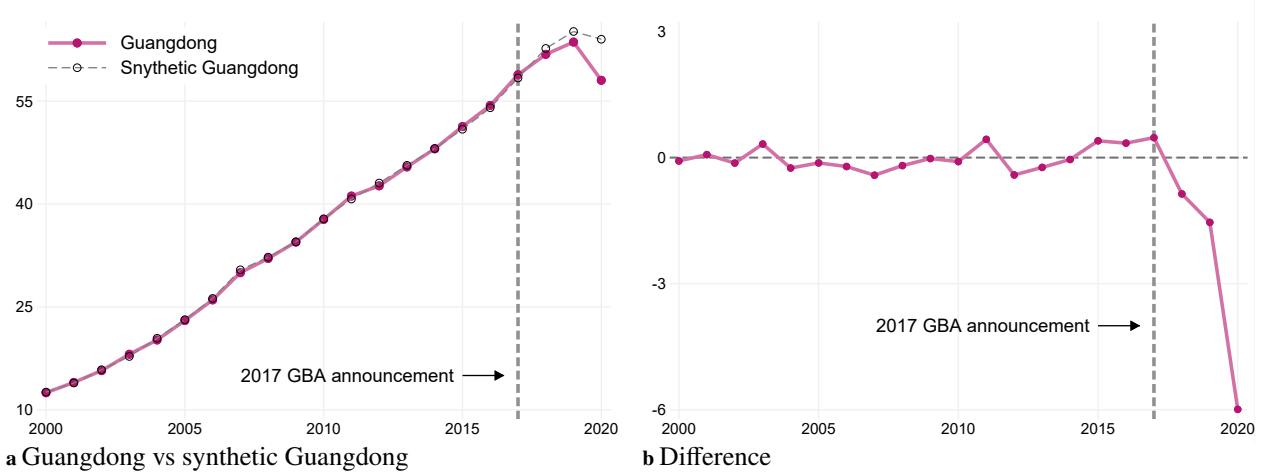


Figure 5. Per Capita GDP of Guangdong and synthetic Guangdong. Panel a shows the path plots for Guangdong vs. synthetic Guangdong (dashed gray line). Synthetic Guangdong is constructed using the synthetic control method where the synthetic counterfactual Guangdong is the weighted average of the other 30 mainland provinces, where weights are chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Panel b shows the difference over time between Guangdong and synthetic Guangdong (Equation (3)). Cumulative differences are show in Figure A9. Dashed vertical line indicates the 2017 announcement of the GBA.

4.2 Inference, Robustness, and Placebo Tests

In this section, we validate our results using a series of sensitivity tests: (i) “placebo-in-place”, (ii) “leave-one-out” analysis, and (iii) “placebo-in-time”, as suggested by [Abadie 2020](#). The first sensitivity analysis, placebo-in-place, also allows us to assess the extent to which the drop in Guangdong’s Per Capita GDP is statistically significant.

[Figure 5 near here]

Our main results for Per Capita GDP reported in [Figure 5](#) are significant to the extent that the observed difference between Guangdong and synthetic Guangdong does not also manifest in the other mainland provinces. We operationalize this intuition as follows. First, we do the placebo-in-place analysis where we iterate over the other 30 mainland provinces (right panel of [Figure 6](#)). In each iteration, one of the donor provinces is the placebo-treated province, and we re-estimate the synthetic control unit for that placebo unit. The difference between the Per Capita GDP of the placebo unit and its synthetic control unit is the estimated placebo treatment effect ($\hat{\tau}_{pt}$).

The full set of placebo results, plus the original results for Guangdong, is reported in [Table 2](#).

Column 5 reports the post-2017 GBA announcement's root mean square prediction error (RMSPE). The post-period RMSPE value is large if the difference between actual and synthetic is large, as defined in [Equation \(3\)](#).

[[Figure 6](#) near here]

We then derive the exact p-value using the distribution of estimated placebo effects, where the p-value for Guangdong is based on how extreme the estimated effect (the post-period RMSPE) for Guangdong is compared to the 30 other placebo units:

$$\text{p-value} = \Pr \left(\left| \bar{\hat{\tau}}_{pt} \right| \geq \left| \bar{\hat{\tau}}_{Gt} \right| \right),$$

where p is one of the 30 donor provinces, $\bar{\hat{\tau}}_{Gt}$ is the mean of the estimated post-period RMSPE for Guangdong, and $\bar{\hat{\tau}}_{pt}$ is the mean of the estimated post-period RMSPE for a donor province j that is not Guangdong.⁹

[[Table 2](#) near here]

Given the distribution of post-period RMSPE for Guangdong and the 30 mainland provinces as placebo units (column 5 of [Table 2](#)), we operationalize the p-value for Guangdong as:¹⁰

$$\text{p-value}_G = \frac{1 + \sum_p^{\mathcal{P}} \mathbb{1} \left(\left| \bar{\hat{\tau}}_{pt} \right| \geq \left| \bar{\hat{\tau}}_{Gt} \right| \right)}{1 + |\mathcal{P}|}. \quad (4)$$

[[Figure 7](#) near here]

⁹Corresponding to [Equation \(3\)](#),

$$\bar{\hat{\tau}}_{Gt} = \frac{\sum_{t=2018}^{2020} \left(Y_{Gt} - \hat{Y}_{Gt} \right)}{3}.$$

¹⁰The p-value for the 30 mainland provinces as placebo units are computed in a similar way by swapping out the G subscript:

$$\text{p-value}_i = \frac{1 + \sum_{j \neq i}^{\mathcal{J}} \mathbb{1} \left(\left| \bar{\hat{\tau}}_{jt} \right| \geq \left| \bar{\hat{\tau}}_{it} \right| \right)}{1 + |\mathcal{J}|}.$$

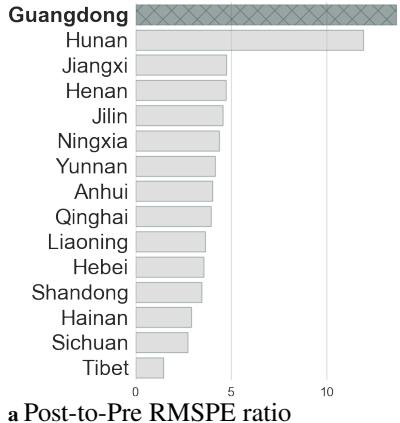
where \mathcal{J} is the set of all 31 mainland provinces including Guangdong.

Table 2. Permutation-based inference for Per Capita GDP

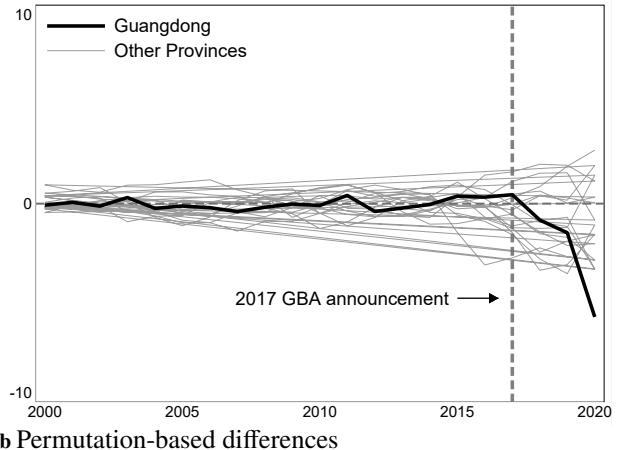
Province	Type	Rank	P-value	Pre	Post	Adjustment for fit quality using RMSPE			
				RMSPE	RMSPE	Post-to-Pre RMSPE	Adjusted Rank	Adjusted P-value	Adjusted Z-score
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Guangdong	Treated	13	0.42	0.26	3.62	13.65	1	0.03	2.14
Hunan	Donor	15	0.48	0.26	3.04	11.92	2	0.06	1.85
Zhejiang	Donor	5	0.16	1.14	9.83	8.61	3	0.10	1.66
Tianjin	Donor	7	0.23	1.37	7.77	5.67	4	0.13	1.52
Gansu	Donor	6	0.19	1.73	8.41	4.85	5	0.16	1.40
Jiangxi	Donor	27	0.87	0.32	1.55	4.77	6	0.19	1.30
Henan	Donor	26	0.84	0.33	1.57	4.73	7	0.23	1.21
Jilin	Donor	16	0.52	0.65	2.95	4.58	8	0.26	1.13
Fujian	Donor	2	0.06	4.51	19.87	4.40	9	0.29	1.06
Ningxia	Donor	19	0.61	0.51	2.23	4.39	10	0.32	0.99
Yunnan	Donor	22	0.71	0.50	2.07	4.17	11	0.35	0.93
Anhui	Donor	28	0.90	0.30	1.21	4.04	12	0.39	0.86
Qinghai	Donor	23	0.74	0.52	2.06	3.94	13	0.42	0.81
Liaoning	Donor	20	0.65	0.61	2.22	3.65	14	0.45	0.75
Hubei	Donor	3	0.10	3.80	13.76	3.62	15	0.48	0.70
Hebei	Donor	18	0.58	0.62	2.23	3.57	16	0.52	0.65
Shandong	Donor	24	0.77	0.57	1.99	3.47	17	0.55	0.60
Beijing	Donor	4	0.13	3.24	10.08	3.12	18	0.58	0.55
Inner Mongolia	Donor	11	0.35	1.36	4.11	3.02	19	0.61	0.51
Hainan	Donor	21	0.68	0.75	2.19	2.93	20	0.65	0.46
Xinjiang	Donor	14	0.45	1.12	3.18	2.84	21	0.68	0.42
Jiangsu	Donor	1	0.03	7.25	19.92	2.75	22	0.71	0.37
Sichuan	Donor	31	1.00	0.12	0.32	2.74	23	0.74	0.33
Shanghai	Donor	8	0.26	2.95	6.90	2.34	24	0.77	0.29
Heilongjiang	Donor	12	0.39	1.67	3.89	2.33	25	0.81	0.25
Guangxi	Donor	9	0.29	2.41	5.32	2.21	26	0.84	0.20
Guizhou	Donor	10	0.32	2.39	4.67	1.96	27	0.87	0.16
Chongqing	Donor	17	0.55	1.50	2.32	1.55	28	0.90	0.12
Tibet	Donor	30	0.97	0.36	0.52	1.45	29	0.94	0.08
Shanxi	Donor	25	0.81	1.65	1.57	0.95	30	0.97	0.04
Shaanxi	Donor	29	0.94	1.43	1.09	0.76	31	1.00	0.00

Table shows the synthetic control method by using each of the 31 mainland provinces as the treated unit. Column 1 indicates that Guangdong is the main treated unit of concern. Column 2 and column 3 are the ranking and p-value based on the highest post-period RMSPE (root mean square prediction error). Column 4 and column 5 are the pre and post-period RMPSE (Equation (3)). Column 6 is the ratio of the post-to-pre RMSPE which adjusts the post-period RMSPE by the pre-period RMSPEs (see Figure 6 and Figure A8). Columns 7–9 are the ranking, p-value and the associated Z-score after the post-period RMSPE is adjusted by fit quality in the pre-period.

As reported in column 3 of Table 2, the p-value for Guangdong is 0.42, which is statistically insignificant. This finding, however, may be an artifact of the placebo units having large post-period RMSPE, as well as large pre-period RMPSE, because of poor fit quality in the estimation



a Post-to-Pre RMSPE ratio



b Permutation-based differences

Figure 6. Permutation-based inference for Per Capita GDP. Left panel shows the distribution of the RMPSE (root mean square prediction error), with the full list in [Table 2](#). Shaded bar is Guangdong. Right panel shows the synthetic control method results by using each of the other 30 mainland provinces in turn as the placebo in gray. The original synthetic control method result for Guangdong from [Figure 5](#) in black. Dashed vertical line indicates the 2017 announcement of the GBA. Both panels retain only mainland provinces with good pre-period fit quality (RMPSE less than the median) for better resolution.

stage ([Equation \(2\)](#) and [Equation \(1\)](#)). Hence, and as recommended by [Abadie 2020](#), we scale the post-period RMSPE by the inverse of the pre-period RMSPE before computing the adjusted p-values:

$$\text{adjusted p-value}_G = \frac{1 + \sum_p^{\mathcal{P}} \mathbb{1} (|\bar{r}_{pt}| \geq |\bar{r}_{Gt}|)}{1 + |\mathcal{P}|}, \quad (5)$$

where the \bar{r} 's are the ratio of the post-period RMSPE (column 6 in [Table 2](#)).¹¹

Based on the adjusted p-value, which accounts for the fit quality, the drop in the Per Capita GDP of Guangdong is now statistically significant at the 5 percent level, as reported in column 8 of [Table 2](#).

This finding confirms that the dip in Guangdong's Per Capita GDP coinciding with the 2017 GBA announcement is unusual compared to China's other mainland provinces. We also report the ratio in the post-to-pre periods RMSPE in the left panel of [Figure 6](#), which shows that the adjusted post-period deviation in actual and predicted Per Capita GDP is largest for Guangdong.

¹¹The RMPSE ratio (column 6 of [Table 2](#)) is computed as:

$$\bar{r}_{jt} = \frac{\bar{\tau}_{jt}}{\sum_{t=2017}^{2020} (Y_{jt} - \hat{Y}_{jt})}.$$

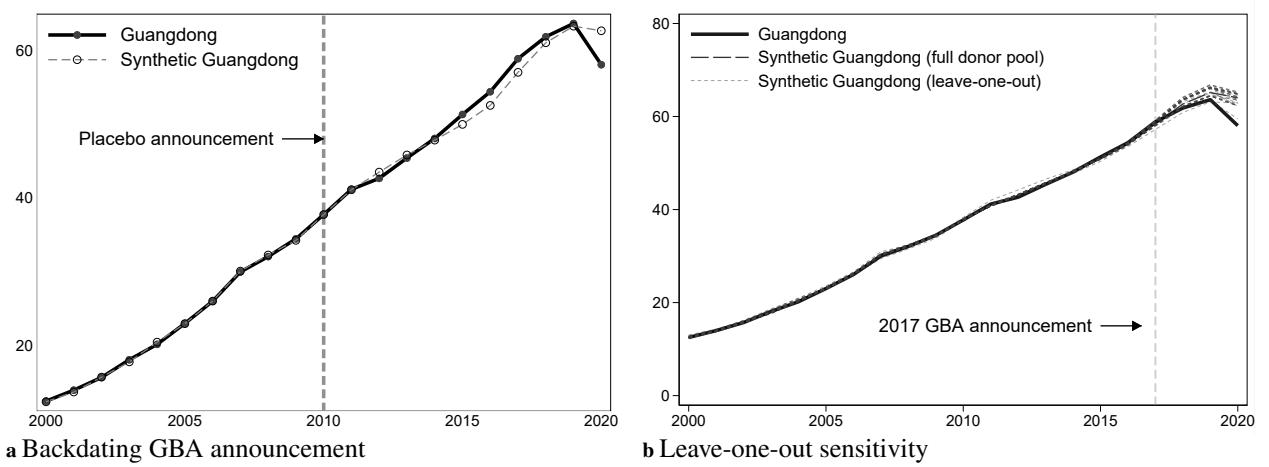


Figure 7. Sensitivity of results for Per Capita GDP. Left panel repeats the synthetic control method to estimate Per Capita GDP for synthetic Guangdong but with the GBA announcement backdated to 2010. Right panel shows the results from a leave-one-out synthetic control estimation where in each of 30 iterations, one of the donor provinces is taken out of the donor pool before estimation. Dashed vertical line indicates the 2017 announcement of the GBA.

We further perform four additional sets of sensitivity tests. First, we perform a leave-one-out analysis where we repeat the synthetic control estimation for Guangdong over 30 iterations. Each iteration drops one of the 30 mainland provinces as a donor province before estimating the path of Per Capita GDP for synthetic Guangdong in the synthetic control estimation. The right panel of Figure 7 reports results from the leave-one-out analysis, with the dotted lines showing the Per Capita GDP of synthetic Guangdong where donor provinces are iteratively dropped from the donor pool. Suppose the exclusion of certain donor provinces greatly affects the path of the estimated synthetic Guangdong. In that case, idiosyncratic macroeconomic shocks or other interventions related to the donor provinces may lead to an artificially large estimated effect in our synthetic control estimation. We find that this is not the case. The Per Capita GDP of the synthetic Guangdong's from the leave-one-out are still able to track Guangdong before 2017 and is higher than Guangdong after the 2017 announcement, especially in the final period in our sample, showing that the main estimated effect of the announcement is robust to the set of donor provinces.¹²

As the second test of sensitivity, we perform a placebo-in-time analysis where we backdate the 2017 GBA announcement to 2010 before repeating the synthetic control estimation with the year 2010 as the placebo GBA announcement year. The left panel of Figure 7 reports the results from

¹²Iterations that fail to converge are omitted from the right panel of Figure 7.

our placebo-in-time analysis and yields three observations. First, synthetic Guangdong estimated using only the period 2000–2010 still closely follows Per Capita GDP of Guangdong, both before and after 2010, where the sample ends for the estimation of the variable and synthetic control weights. Second, the gap in Per Capita GDP between Guangdong and synthetic Guangdong still appears after the 2017 announcement and is most distinct in the final year of our sample period, even when the synthetic control estimation procedure no longer has information on the 2017 announcement date. Finally, the absence of a gap in Per Capita GDP between Guangdong and synthetic Guangdong before 2017, even when the synthetic control estimation does not enforce a match, mitigates concerns of anticipatory effects.

Third, we extend the set of predictor variables with five additional variables: (i) illiteracy rate (percentage of population aged 15 and above), (ii) tertiary institutions (percentage of institutions), (iii) tourist arrivals (normalized by GDP), (iv) realized foreign direct investments (excluding regional investment and normalized by GDP), and (v) realized investments (normalized by GDP). We then re-estimate the synthetic control unit for both Guangdong and the 30 other mainland provinces to obtain the adjusted p-value for Guangdong, which turns out to be unchanged ([Figure A10](#)). Guangdong’s Per Capita GDP after the announcement, as measured by the adjusted post-RMSPE, is still unusually low relative to the other provinces. Finally, we repeat the synthetic control estimation with a lower threshold margin for constraint violation in the constrained quadratic optimization, and the conclusions are similar ([Figure A11](#)).

5 Potential Effect on Population Density

5.1 *Results*

This section shows that the increase in population density can explain the decline in Per Capita GDP for Guangdong.

We repeat the synthetic control method described in [Section 2](#) and [Section 4](#), this time with Population density as the outcome variable. The set of predictor variables and the pre-intervention

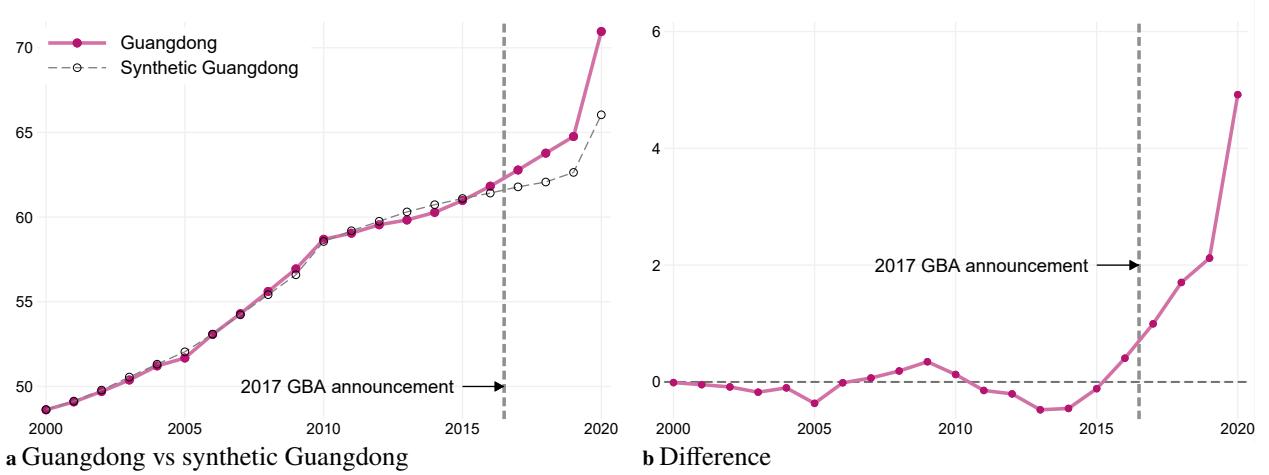


Figure 8. Population density of Guangdong and synthetic Guangdong. Panel a shows the path plots for Guangdong vs. synthetic Guangdong (dashed gray line). Synthetic Guangdong is constructed using the synthetic control method where the synthetic counterfactual Guangdong is the weighted average of the other 30 mainland provinces, where weights are chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Panel b shows the difference over time between Guangdong and synthetic Guangdong (Equation (3)). Dashed vertical line indicates the 2017 announcement of the GBA.

estimation period are otherwise the same. [Figure 8](#) the main results and the [Supplementary Materials](#) reports the full set of estimation results. Panel a of [Figure 8](#) shows that the Population density of synthetic Guangdong tracks the Population density of Guangdong well until 2017 before diverging. More specifically, panel b of [Figure 8](#) suggests that, before the announcement, the difference between Guangdong and synthetic Guangdong is zero. After the announcement, the graph shows an upward trend.

[[Figure 8](#) near here]

The synthetic control results confirm Guangdong’s population density increased after the 2017 GBA announcement. Further, we find quantitative evidence that the increase in Guangdong’s population density coincides with the GBA announcement in a way that is not reflected in the other mainland provinces.

5.2 Inference, Robustness, and Placebo Tests

Similar to [Section 4.2](#), we validate our results for the drastic increase in Guangdong’s Population density. First, we do the placebo-in-place analysis (reported in [Table A4](#) and [Figures A16 to A17](#)).

Here, we find that Guangdong indeed had a population growth that is different from the rest of the mainland provinces. In particular, the sharp increase in Guangdong’s population can be observed in 2019 ([Figure 2](#) and [Figure A16](#)).

Second, we perform a leave-one-out analysis and get similar results ([Figure A15](#)). Third, we perform the placebo-in-time analysis which shows that even when we backdate the GBA announcement to 2010, and the estimation has no information on when the GBA announcement happened, a divergence still occurs after 2017. Fourth, we extend the set of predictors to the same five additional variables in [Section 4.2](#) and get same results for Guangdong’s Population density ([Figure A18](#)). Finally, we lower the threshold margin for constraint violation in the synthetic control estimation and get similar findings ([Figure A19](#)).

Overall, we find that the conclusion of a sharp increase in Guangdong’s Population density leading to a recorded drop in Per Capita GDP using the synthetic control estimation is robust to a variety of sensitivity and placebo tests.

6 Discussion

6.1 *Labour Productivity*

[[Figure 9](#) near here]

Since we find a decline in Guangdong’s Per Capita GDP and that the rise in population growth for Guangdong is a key contributing factor, a natural question is whether we would find similar changes in productivity. We use GDP in 1m yuan per 1,000 workers to track the productivity of the provinces. Repeating the same synthetic control estimation in [Sections 4](#) to [5](#), we focus on Guangdong’s productivity and examine how it changes after the 2017 GBA announcement.

[Figure 9](#) shows a decline in Guangdong’s productivity relative to trend in productivity for synthetic Guangdong after 2018. This decline, at least in our sample period ending 2020, is persistent, as shown in the right panel of [Figure 9](#). From our permutation-based inference, we again

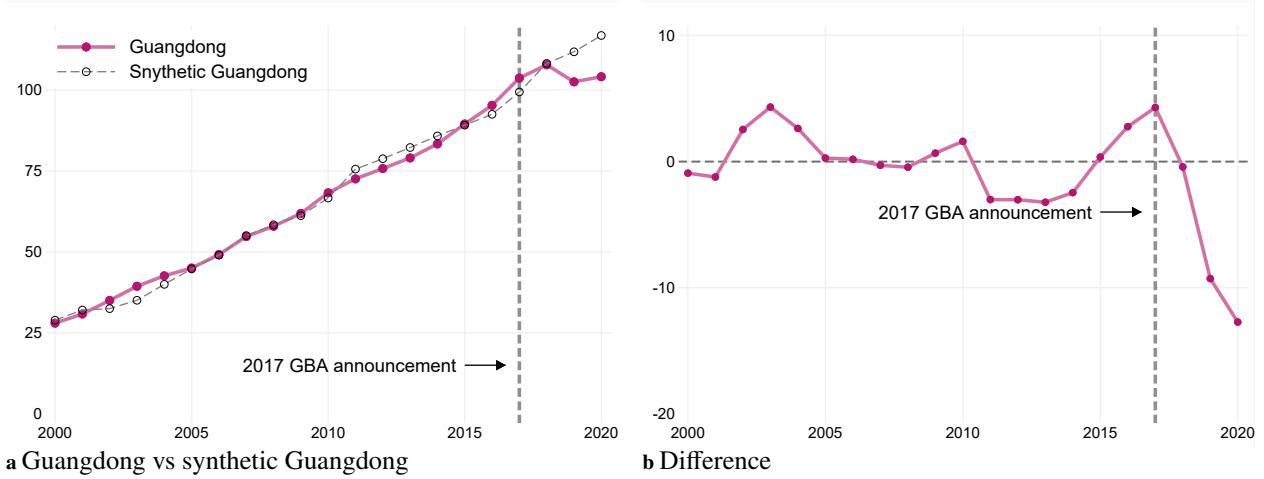


Figure 9. Labour Productivity of Guangdong and synthetic Guangdong. Panel a shows the path plots for Guangdong vs. synthetic Guangdong (dashed gray line). Synthetic Guangdong is constructed using the synthetic control method where the synthetic counterfactual Guangdong is the weighted average of the other 30 mainland provinces, where weights are chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Panel b shows the difference over time between Guangdong and synthetic Guangdong (Equation (3)). Dashed vertical line indicates the 2017 announcement of the GBA.

confirm that this decline in productivity is unusual compared to the other provinces.¹³

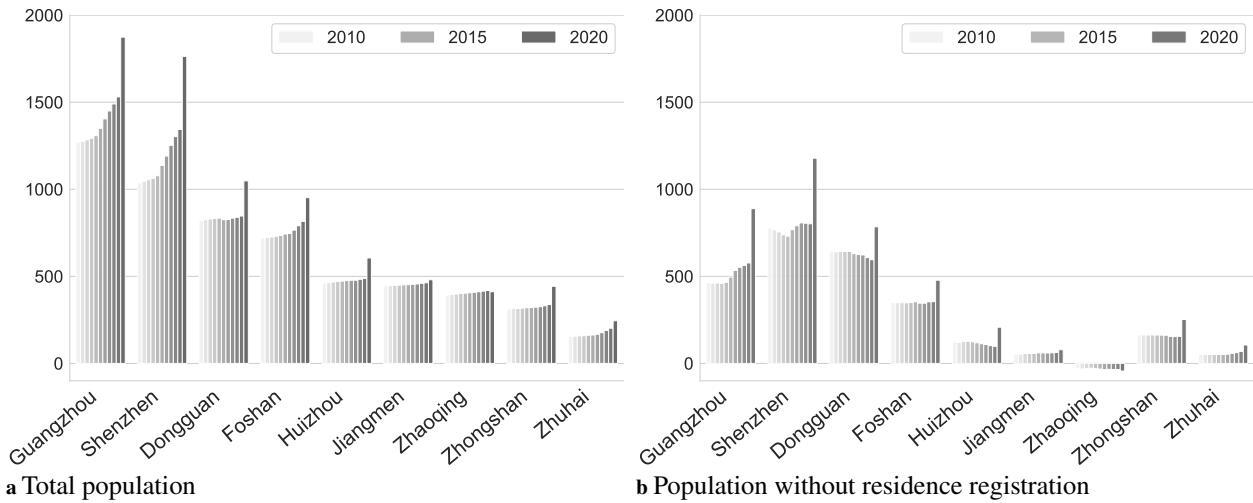
The decline in productivity may be attributed to the significant rise in employment in relatively lower-productivity sectors, such as construction, wholesale, and retail.¹⁴ Based on the National Bureau of Statistics, the number of workers in these three sectors has increased from 15.4 million yuan in 2015 to 23.1 million yuan in 2019.¹⁵ On the other hand, even though employment in the ICT and finance sector doubled (from 0.8 million yuan to 1.7 million yuan) during the same period, the lower employment level could not significantly increase overall productivity. In the next section, we focus on the talent attraction policies.

The influx of talents should increase Per Capita GDP in general, which is observed in the trends and our findings. This does not necessarily imply that the GBA policy broadly, nor the attraction of talents, did not work in spur the economy in Guangdong. Instead, we conjecture that two reasons

¹³All other provinces with larger deviations from their counterfactual have an increase instead of a decrease in productivity (Table A6). Hence, for provinces with a decline in productivity, Guangdong is the one with the largest decline relative to the counterfactual. This is best seen in Figure A24 of the Supplementary Materials which traces out the pre- and post-period difference for each province and its synthetic control unit.

¹⁴According to the Fourteenth Five Year Plan of Guangdong Province Construction Sector (2021), the productivity level of the construction sector in 2020 is around RMB 500,000 per capita, lower than that of Hubei Province, Shanghai, and Beijing (see <http://zfcxjst.gd.gov.cn/attachment/0/425/425095/3309876.pdf>.)

¹⁵The employment number includes urban employed persons, employed persons in urban private enterprises, and self-employed individuals in urban areas.



a Total population

b Population without residence registration

Figure 10. Population ('0,000) in Pearl River Delta cities (2010–2020). City population without residence registration is the total population minus the population with residence registration by year-end. The last bar for each city is for the year 2020.

for the observation that Per Capita GDP seems to decline even as the region attracted more talent. One is that the rise in employment is concentrated in relatively lower productivity sectors such as construction, wholesale, and retail. From the National Bureau of Statistics, the three sectors had an increase of 7.7 million workers between 2015 and 2019. The second is that our results are only short-run. It may take some time for the economy to recalibrate before the fruits of the new talents can be reaped in the medium run. We discuss and contextualize more with the population-related policies in the following section.

[Figure 10 near here]

6.2 Population and Talent Attraction in the GBA

Our results above suggest that the decline, and the timing of the decline, in Guangdong's Per Capita GDP can be attributable to an unusually large increase in Guangdong's population (Figure 5 and Figure 8).

Figure 10 shows the city-level population data for the nine Pearl River Delta cities in Guangdong. Many of these cities experience a distinct jump in population corresponding to Guangdong's province-wide population jump in 2019–2020.¹⁶ In particular, Guangzhou, Shenzhen, Dongguan,

¹⁶Total population includes all the people who live in the city for more than six months. The residential population

Foshan, Huizhou, and Zhongshan all have a jump in population from 2010–2020, which suggests structural changes (e.g., [Lee and Lin 2020](#)). These jumps in the population likely relate, at least partially, to labour inflow since Guangdong and the nine Pearl River Delta cities all have policies related to talent recruitment for the GBA. Below, we walk through some key talent recruitment policies of the GBA.

One key policy can be traced back to 2017 when the Guangdong government issued the Opinions on Deepening the Reform of the Institutional Mechanism for Talent Development which aims to establish a globally competitive talent system. It includes 16 entry policies for foreign talents, raising the R&D expenditure for firms, and promoting entrepreneurship for science and technology start-ups. The government provides housing subsidies up to RMB 3.5 million, RMB 2.5 million, and RMB 1.5 million, respectively, for top talents, leading national talents, and leading provincial talents.¹⁷ On 1st January 2019, The Regulations of the Guangdong Talent Development took effect, which covered talent training and development, talent attraction, talent evaluation and incentivisation, and talent services and guarantees.¹⁸

There are three sources for talent: 1) other Mainland China provinces, 2) two SARs – Hong Kong and Macau, and 3) abroad. The relaxation of *hukou* or household registration status policy is the major initiative to recruit talents from other provinces. The benefits of residence (i.e., with local hukou) include healthcare, education for children, and permits to purchase houses. Take Guangzhou, the capital city, as an example. In 2017, people with bachelor's degrees needed to continuously contribute to social insurance for at least one year, whereas the requirement was reduced to six months in 2019. This requirement is also removed for people with a master's degree and above in 2019.¹⁹ Moreover, starting from 2020, talents with bachelor's degrees from “Double First-Class” universities can get their hukou once they start paying for social insurance.²⁰

includes people with Hukou registered in the city. Thus, the population without residence registration is not equal to the migrant population. But we can use it as a proxy for a city's population inflow.

¹⁷Guangdong issued the implementation of the reform of the talent development system mechanism. See http://www.gd.gov.cn/zwgk/zcjd/snzcjd/content/post_76542.html.

¹⁸See Guangdong Provincial Talent Development Regulations at http://gdstc.gd.gov.cn/kjzx_n/gdkj_n/content/post_2696003.html.

¹⁹See http://www.gz.gov.cn/zwgk/zcjd/zcjd/content/post_2854375.html.

²⁰See the Notice of Human Resources and Social Security Bureau of Guangzhou on relaxing the social insurance

For foreign talents, Guangdong issued two Excellent Talent cards (Excellent Talent Card A and Excellent Talent Card B) to attract workers from overseas to work in Guangdong. Both types of cardholders will enjoy housing guarantees, education for children, social insurance, healthcare, long period of stay, and multiple entries. They will have advantages in applying for China's permanent residence. To qualify for the Guangdong Excellent Talent Card A, one has to have received prestigious scientific awards and prizes in academia, be a senior executive of Fortune Global 500 companies, etc. To qualify for Talent Card B, one has to be selected for the National "Thousand Talents Plan" or be a post-doctoral from a top 200 university worldwide, etc.²¹ Moreover, starting from 2019, foreign talents working in the Greater Bay Area will receive a certain amount of income tax subsidy from the local government.²²

For Hong Kong and Macao residents, in addition to all the preferential policies for foreign talents, they also benefit from policies that give preferential treatment relating to law, healthcare, and tourism in 2019.²³ Furthermore, the nine Pearl River Delta cities also have plans to set up specific youth innovation entrepreneurship bases to attract younger individuals from Hong Kong and Macao.²⁴ From 2021, the start-ups in these bases will be provided with rental subsidies of up to 6,000 RMB per year for up to 3 years and a one-time start-up subsidy of 10,000 RMB. Eligible young people from Hong Kong and Macao can apply for up to 5 million yuan loans and get discounted interest rates.²⁵

While we scan a range of GBA policies in Guangdong that potentially explain the influx of population in Guangdong, we believe that most of these account for only a small proportion of

eligibility period for talents with bachelor's degrees from "Double First-class" universities at http://rsj.gz.gov.cn/ywzt/rcgz/rcyjrh/tzgg/content/post_6993948.html.

²¹See Notice of the People's Government of Guangdong Province on Implementation Measures of Guangdong Excellent Talent Card at http://hrss.zs.gov.cn/zcfg/pxjy/content/post_1382027.html.

²²See Notice on The Implementation of The Individual Income Tax Preferential Policies in the Guangdong-Hong Kong-Macao Greater Bay Area at http://czt.gd.gov.cn/czfg/content/post_2519383.html?from=groupmessage&isappinstalled=0.

²³See http://hrss.gd.gov.cn/zwgk/gsgg/content/post_2711821.html.

²⁴See Implementation Plan on Strengthening the Construction of Youth Innovation and Entrepreneurship Bases for Hong Kong and Macao Young Talents at http://www.gd.gov.cn/zwgk/wjk/qbj/yfh/content/post_2469478.html.

²⁵See Opinions on promoting high-quality development of Youth Innovation and Entrepreneurship Bases for Hong Kong and Macao Young Talents at https://www.cnbayarea.org.cn/policy/policy%20release/policies/content/post_319721.html.

Guangdong's growth in population. Our conjecture is that while the GBA has plans and policies for foreign talent growth, most of the Guangdong's population growth still originates from the other mainland provinces. The fall in population in certain provinces in the same year where we observe Guangdong population spike (see the [Provincial population](#) section in the [Supplementary Materials](#)) lends credence to our conjecture.

7 Conclusion

We conduct a quantitative case study of Guangdong as the epicentre of the newly developed GBA in China. Our results suggest that Guangdong's economic performance has declined since the 2017 GBA announcement in a way that is unusual compared to the other mainland provinces. In particular, we trace the decline of Guangdong's Per Capita GDP to an influx in population potentially related to population and talent attraction policies in the GBA. In ongoing studies, we aim to unpeel heterogeneities in demographic shifts occurring in the GBA, as well as studying more closely the patterns and origin of talent migration into the core GBA cities.

Based on our analyses, we stress the observation that basic economic indicators, such as Per Capita GDP and labour productivity, only declined because of large and dramatic demographic shifts. We learn that, had we not systematically expose such shifts in population for key cities of Guangdong, interpreting a negative impact of the GBA on the Guangdong economy would have been a spurious interpretation. Fundamental economic performance indicators may only show upward trends once the sectors have recalibrated from the large demographic shifts, and the new workforce is assimilated into the economy. Such shifts imply that any early evaluation of urbanization or regionalization plans needs to account for such demographic changes to avoid spurious interpretations of policy efficacy.

References

Abadie, Alberto. 2020. “Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects.” *Journal of Economic Literature* 59 (2): 391–425. [10.1257/jel.20191450](https://doi.org/10.1257/jel.20191450).

Abadie, Alberto, Alexis Diamond, and Hainmueller Jens. 2010. “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program.” *Journal of the American Statistical Association* 105 (490): 493–505. [10.1198/jasa.2009.ap08746](https://doi.org/10.1198/jasa.2009.ap08746).

Abadie, Alberto, and Javier Gardeazabal. 2003. “The Economic Costs of Conflict: A Case Study of the Basque Country.” *American Economic Review* 93 (1): 113. [10.1257/aer.101.6.2590](https://doi.org/10.1257/aer.101.6.2590).

Bin, WANG, and YANG Lu. 2020. “World-Class Bay Area Experience and Coordinated Development of the Guangdong-Hong Kong-Macao Greater Bay Area.”

Chow, Man Kong, Jingbo Hua, and Wing Lok Hung. 2020. “Tertiary Education and Innovation in the Greater Bay Area.” *Asian Education and Development Studies* 9 325–336. [10.1108/AEDS-02-2019-0048](https://doi.org/10.1108/AEDS-02-2019-0048).

Ding, Hong-chao, Min-rong Lian, Xue-ying Chen, Jia-ming Liu, Zu-chang Zhong, Yu-fan Zhang, and Min-yao Zhou. 2019. “Research on the Correlation of Port Logistics and Regional Economic Growth Base on Gray Relational Analysis Method.” *Concurrency and Computation: Practice and Experience* 31 e4744. [10.1002/cpe.4744](https://doi.org/10.1002/cpe.4744).

Hu, Shanshan, and Hyung-Ho Kim. 2022. “Research on Urban Innovation Efficiency of Guangdong-Hong Kong-Macao Greater Bay Area Based on DEA-Malmquist Model.” *Annals of Operations Research* 1–14. [10.1007/S10479-022-04577-8/FIGURES/2](https://doi.org/10.1007/S10479-022-04577-8/FIGURES/2).

Hui, Eddie C.M., Xun Li, Tingting Chen, and Wei Lang. 2020. “Deciphering the Spatial Structure of China’s Megacity Region: A New Bay Area—the Guangdong-Hong Kong-Macao Greater Bay Area in the Making.” *Cities* 105 102168. [10.1016/J.CITIES.2018.10.011](https://doi.org/10.1016/J.CITIES.2018.10.011).

Lee, Ivan, and Regina Fang Ying Lin. 2020. “Economic Complexity of the City Cluster in Guangdong-Hong Kong-Macao Greater Bay Area, China.” *Sustainability (Switzerland)* 12. [10.3390/SU12145639](https://doi.org/10.3390/SU12145639).

Liu, Yixin, and Miao Zhang. 2021. “Research on Financial Management of Guangdong-Hong Kong-Macao Greater Bay Area Based on LS-SVM Algorithm and Multi-Model Fusion.” *Annals of Operations Research* 1–19. [10.1007/S10479-021-04398-1/TABLES/5](https://doi.org/10.1007/S10479-021-04398-1/TABLES/5).

Loo, P. Y. Becky, and WANG Bo. 2018. “The Importance of Integrated Transport in Fostering the Formation of the Guangdong-Hong Kong-Macao Greater Bay Area.” *Progress in Geography* 37 (12): 1623. [10.18306/dlkxjz.2018.12.004](https://doi.org/10.18306/dlkxjz.2018.12.004).

Ni, Wai, Shihua Zhou, and Zhiyu Wei. 2020. “Research on the Development of Economic Integration in Bay Area: Based on the Gungdong-Hong Kong-Macao Greater Bay Area.”

Sheng, Yanwen, Jinli Zhao, Xuebo Zhang, Jinping Song, and Yi Miao. 2019. “Innovation Efficiency and Spatial Spillover in Urban Agglomerations: A Case of the Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Pearl River Delta.” *Growth and Change* 50 1280–1310. [10.1111/GROW.12329](https://doi.org/10.1111/GROW.12329).

Tang, Rui. 2021. “The Impact of Integration Policies on Tourism Industry Convergence in the Yangtze River Delta: Theoretical Mechanism and Empirical Test.” *Letters in Spatial and Resource Sciences* 14 141–167. [10.1007/S12076-021-00270-9/TABLES/9](https://doi.org/10.1007/S12076-021-00270-9/TABLES/9).

Wen, Yuyuan. 2014. “The Spillover Effect of FDI and Its Impact on Productivity in High Economic Output Regions: A Comparative Analysis of the Yangtze River Delta and the Pearl River Delta, China.” *Papers in Regional Science* 93 341–365. [10.1111/PIRS.12086](https://doi.org/10.1111/PIRS.12086).

Weng, Hongtao, Jia Kou, and Qinglong Shao. 2020. “Evaluation of Urban Comprehensive Carrying Capacity in the Guangdong–Hong Kong–Macao Greater Bay Area Based on Regional Collaboration.” *Environmental Science and Pollution Research* 27 20025–20036. [10.1007/S11356-020-08517-6/TABLES/5](https://doi.org/10.1007/S11356-020-08517-6/TABLES/5).

Xu, Hengzhou, and Man Jiao. 2021. “City Size, Industrial Structure and Urbanization Quality—A Case Study of the Yangtze River Delta Urban Agglomeration in China.” *Land Use Policy* 111 105735. [10.1016/J.LANDUSEPOL.2021.105735](https://doi.org/10.1016/J.LANDUSEPOL.2021.105735).

You, Xiaojun, Qixiang Li, Kyle M. Monahan, Fei Fan, Haiqian Ke, and Na Hong. 2022. “Can Collaborative Innovation Constrain Ecological Footprint? Empirical Evidence from Guangdong–Hong Kong–Macao Greater Bay Area, China.” 3. [10.1007/s11356-022-19648-3](https://doi.org/10.1007/s11356-022-19648-3).

Zhang, Chi, Zhang Xuyao, and Zhou Jingwei. 2022. “Development Potential and New Opportunities for Node Cities in GBA: A Case Study of Foshan.” <https://lkyspp.nus.edu.sg/docs/default-source/aci/acirp202204.pdf>.

Supplementary Materials for “Income and Productivity Trends of Guangdong: a Data-driven Case Study of the Greater Bay Area” (Jingwei Li, Lucas Shen, Xuyao Zhang, May 2023)

Table of Contents

- A. Data Description and Population Data Source
- B. Supplementary Notes on Guangdong and the GBA
- C. Provincial Population
- D. Additional Tables and Figures for Per Capita GDP
- E. Additional Tables and Figures for Population Density
- F. Additional Tables and Figures for Productivity

A. Data Description and Population Data Source

Table A1. Notes on annual province population data

Year	Type	Source	Additional/misellaneous notes
2000	Decennial census	NBS accessed on 21 Dec 2021	—
2001	Revised	NBS accessed on 21 Dec 2021	—
2002	Revised	NBS accessed on 21 Dec 2021	—
2003	Revised	NBS accessed on 21 Dec 2021	—
2004	Revised	NBS accessed on 21 Dec 2021	—
2005	Revised	NBS accessed on 21 Dec 2021	—
2006	Revised	NBS accessed on 21 Dec 2021	—
2007	Revised	NBS accessed on 21 Dec 2021	—
2008	Revised	NBS accessed on 21 Dec 2021	—
2009	Revised	NBS accessed on 21 Dec 2021	—
2010	Decennial census	NBS accessed on 21 Dec 2021	—
2011	Unrevised	Yearbook 2012	Hunan's data is from yearbook2021
2012	Unrevised	Yearbook 2013	Hunan's data is from yearbook2021
2013	Unrevised	Yearbook 2014	Hunan's data is from yearbook2021
2014	Unrevised	Yearbook 2015	Hunan's data is from yearbook2021
2015	Unrevised	Yearbook 2016	Hunan's data is from yearbook2021
2016	Unrevised	Yearbook 2017	Hunan's data is from yearbook2021
2017	Unrevised	Yearbook 2018	Hunan's data is from yearbook2021
2018	Unrevised	Yearbook 2019	Hunan's data is from yearbook2021
2019	Unrevised	Yearbook 2020	Hunan's data is from yearbook2021; Hebei Yearbook2020 hasn't release, the data is from NBS accessed on 4 Oct 2021
2020	Decennial census	NBS accessed on 14 Dec 2021	

NBS refers to the National Bureau of Statistics of China.

Table A2. Data description and summary

Variable	Description	N	Mean	Std. Dev.	Min.	Max.
Primary industry	Primary industry value over GDP (both in Yuan, Constant Prices at 2000)	651	0.12	0.07	0.00	0.36
Secondary industry	Secondary industry value over GDP (both in Yuan, Constant Prices at 2000)	651	0.46	0.12	0.16	0.85
Tertiary industry	Tertiary industry value over GDP (both in Yuan, Constant Prices at 2000)	651	0.46	0.09	0.29	0.88
Household income per capita	Per Capita Annual Urban Household Income (Yuan, Constant Prices at 2000)	651	16,053.16	8,217.72	4,745.27	51,682.00
Visitors	International visitor arrivals (person-times) over GDP (Yuan, Constant Prices at 2000)	620	0.01	0.01	0.00	0.12
Unemployment rate	Unemployment rate in urban areas (Percentage)	616	3.50	0.72	0.80	6.50
Annual wage	Average Annual Wage of Staff and Workers (Yuan, Constant Prices at 2000)	620	29,552.82	17,224.01	6,918.00	115,966.00
Illiteracy rate	Population aged 15 and above who are illiterate (Percentage)	619	8.29	7.31	1.23	54.86
Tertiary institutions	Tertiary institutions out of total institutions (Percentage)	620	0.99	1.04	0.06	5.44
Tourist arrivals	Regional tourist arrivals (person-times) over GDP (Yuan, Constant Prices at 2000)	583	1.80	1.20	0.03	10.54
FDI	Realized foreign direct investment (excl. regional investment) over GDP (both in Yuan, Constant Prices at 2000)	523	0.04	0.04	0.00	0.26
Investments	Realized Regional Investment over GDP (both in Yuan, Constant Prices at 2000)	492	1.30	0.63	0.42	7.72
Per Capita GDP	GDP (Yuan, Constant Prices at 2000) over population at year-end	651	25.60	19.10	2.74	114.40
Population density	Persons per Square kilometer	651	42.47	62.38	0.21	394.92
Productivity (labor)	GDP (Yuan, Constant Prices at 2000) over employment at year-end	648	46.15	33.19	5.03	195.97

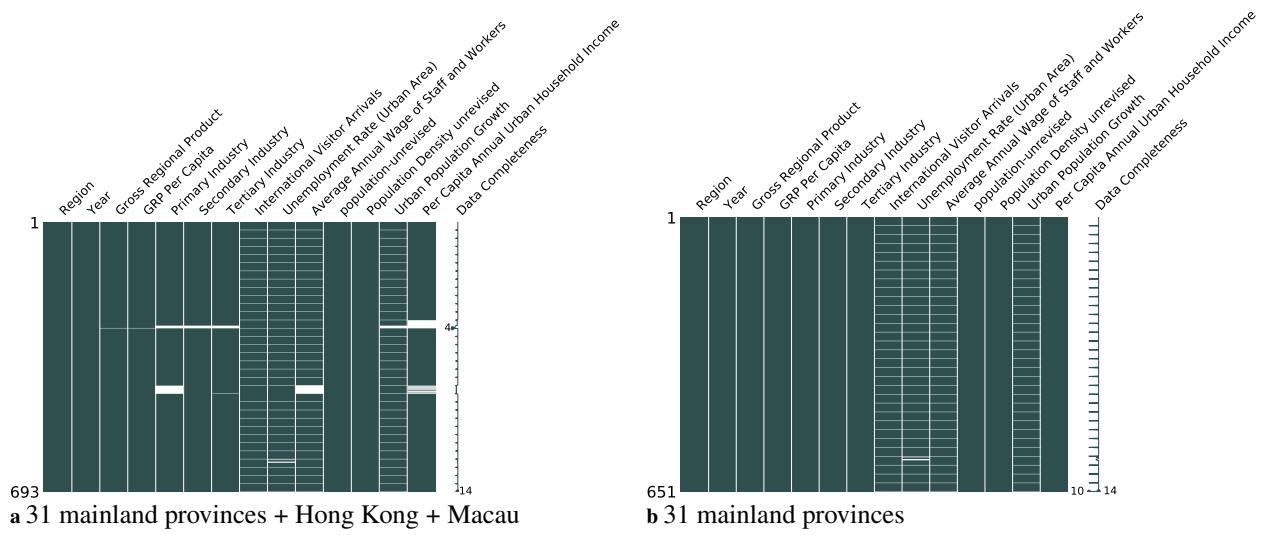


Figure A1. Province-year panel completeness. Each row is a province-year observation and each column is a variable. Holes indicate missing data.

B. Supplementary Notes on Guangdong and the GBA

In 2017, the Framework Agreement on Deepening Guangdong-Hong Kong-Macao Co-operation in the Development of the Bay Area (GBA) was signed, prioritizing the cooperation in infrastructure connectivity, market integration, technology, and innovation within the region. In 2019, the Outline Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area was further announced, mapping out the whole development plan for the GBA. The Outline sets two stages of objectives for 2022 and 2035.

Overall, the GBA aims to build on the economic potential of the Pearl River Delta (PRD) region, which includes the nine Guangdong municipalities: Guangzhou, Shenzhen, Zhuhai, Foshan, Dongguan, Zhongshan, Jiangmen, Zhaoqing, and Huizhou. Since the 1990s, the PRD municipalities have been a leader in China's regional integration and economic development. The GBA, which includes the nine PRD municipalities in Guangdong together with Hong Kong and Macao, aims to become an international innovation hub. This region will leverage on the existing developments in Guangdong, which has attracted leading science and technology enterprises, such as Tencent and Huawei, and abundant talents to boost its R&D capacity. Several platforms and innovation parks have also been established, including the Nansha Guangdong-Hong Kong Cooperation Park, Sino-

Singapore Guangzhou Knowledge City, Artificial Intelligence and Digital Economy Pilot Zone in Guangzhou, Dongguan Binhaiwan Bay Area, Sanlong Bay High-end innovation Cluster Zone in Foshan, and Zhaoqing New Town (http://dfz.gd.gov.cn/sqyl/gmjj/content/post_3266899.html).

While Guangdong is geographically larger than these nine municipalities, in our main analyses (Sections 4 to 6.1), we focus on the Guangdong as the unit of intervention. This approach is because while Guangdong has 21 municipalities, the nine PRD municipalities collectively contribute 80.5% of GDP in the Guangdong province and include 62.0% of the provincial population (based on 2020 data from the Guangdong Statistical Yearbook). Moreover, given that one key pillar of the GBA is innovation, six out of nine PRD municipalities cities—Guangzhou, Shenzhen, Dongguan, Huizhou, Zhuhai, and Zhongshan—took up 75.5% of the number of patents certified in Guangdong in 2020. Hence, whatever gain in economic benefits from the development of the GBA should be reflected in that of the Guangdong province.

We do not focus on the Hong Kong and Macao Special Administrative Regions in our analyses because (i) they have their own economic system, (ii) including the two Special Administrative Regions to Guangdong will likely inflate measures of economic performance for Guangdong and make it less likely that we have a convex combination in the synthetic control estimation, (iii) the two Special Administrative Regions are part of the intervened regions and should not be included in the donor pool when the unit of intervention for the GBA is Guangdong, and (iv) the two Special Administrative Regions have poorer data completeness (Figure A1). However, including the Hong Kong and Macao Special Administrative Regions in the donor does not affect the eventual optimal weights (untabulated).

C. Provincial Population

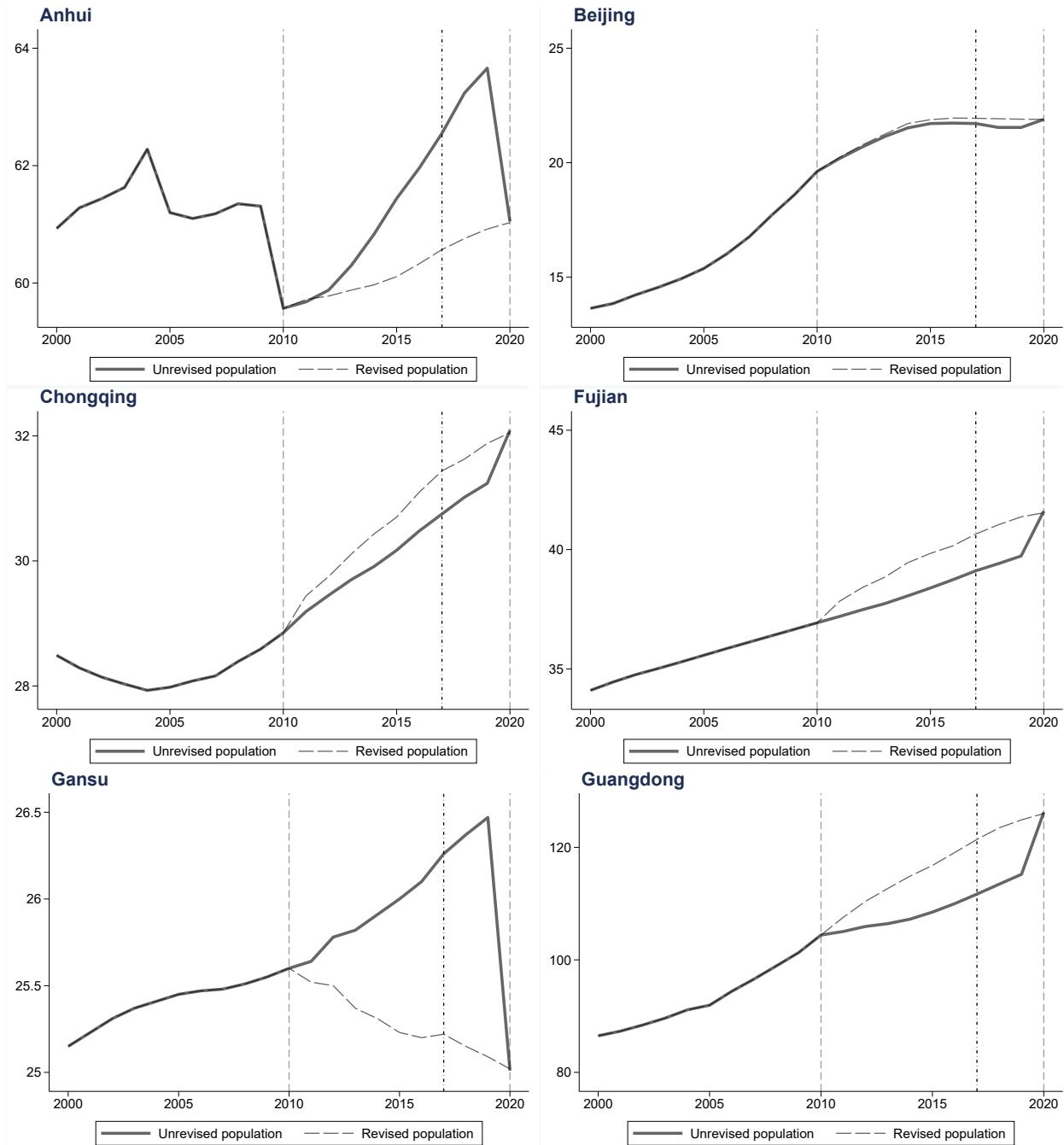


Figure A2. Population path plots for selected provinces [Part 1]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

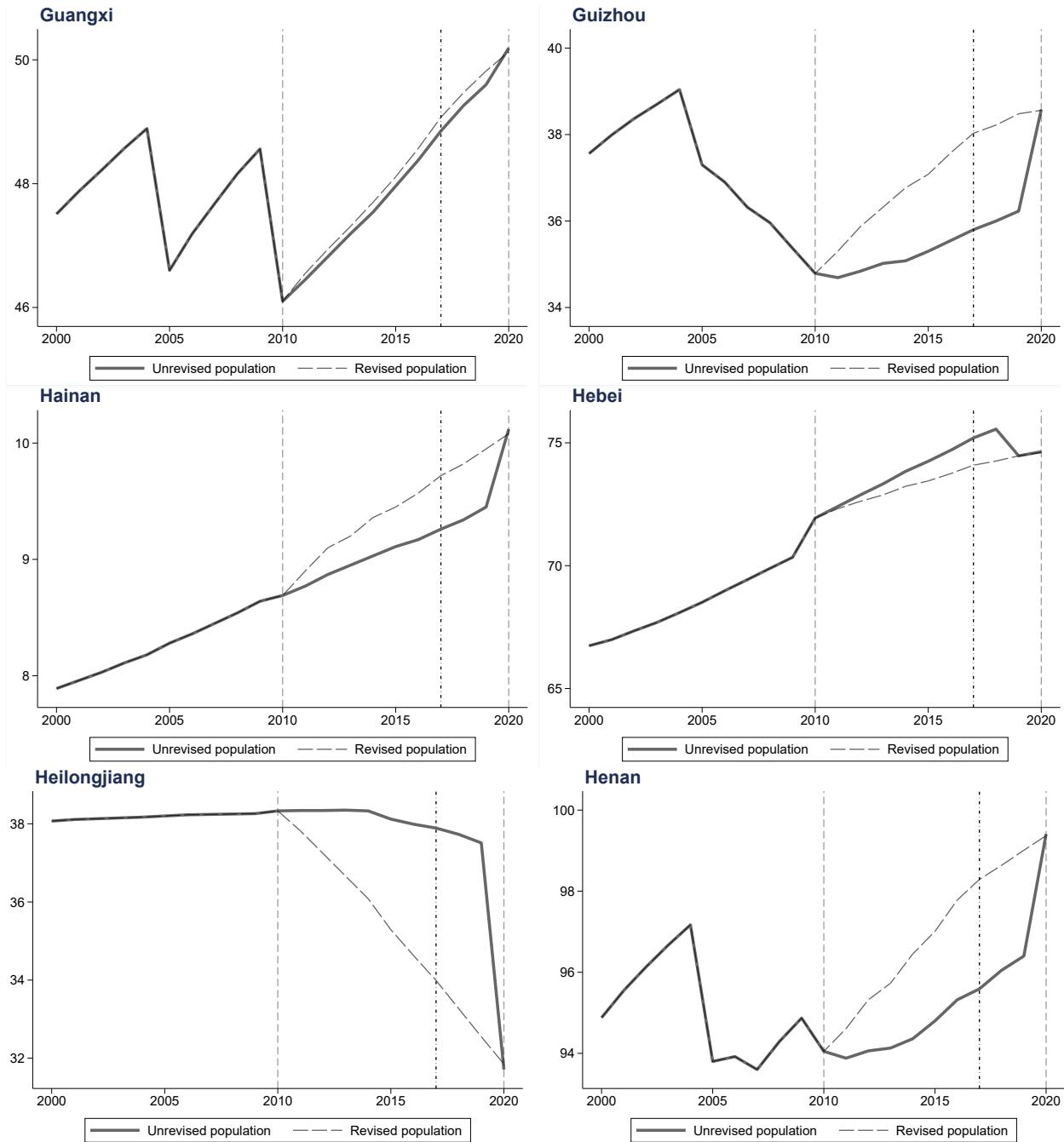


Figure A3. Population path plots for selected provinces [Part 2]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

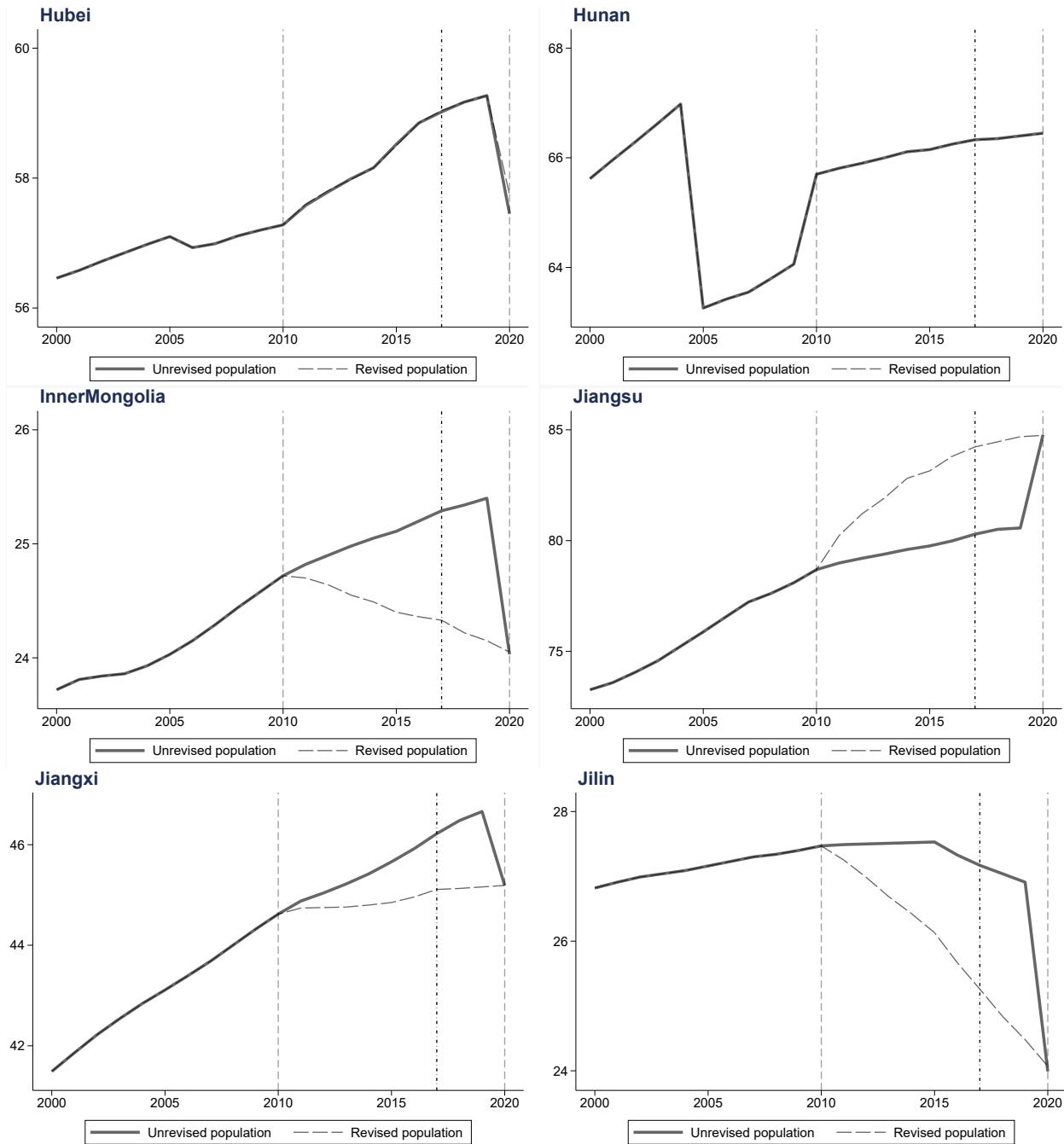


Figure A4. Population path plots for selected provinces [Part 3]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

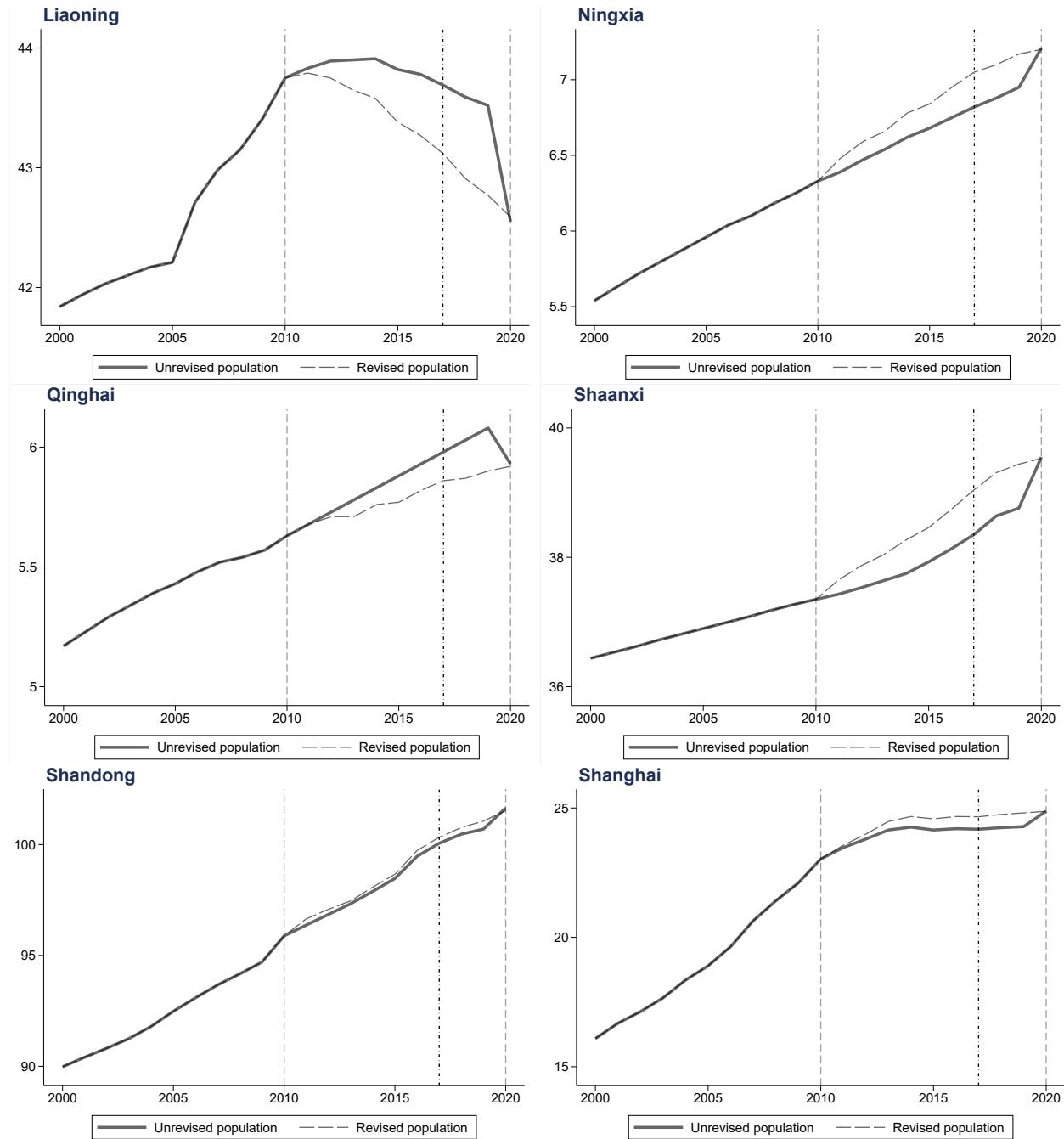


Figure A5. Population path plots for selected provinces [Part 4]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

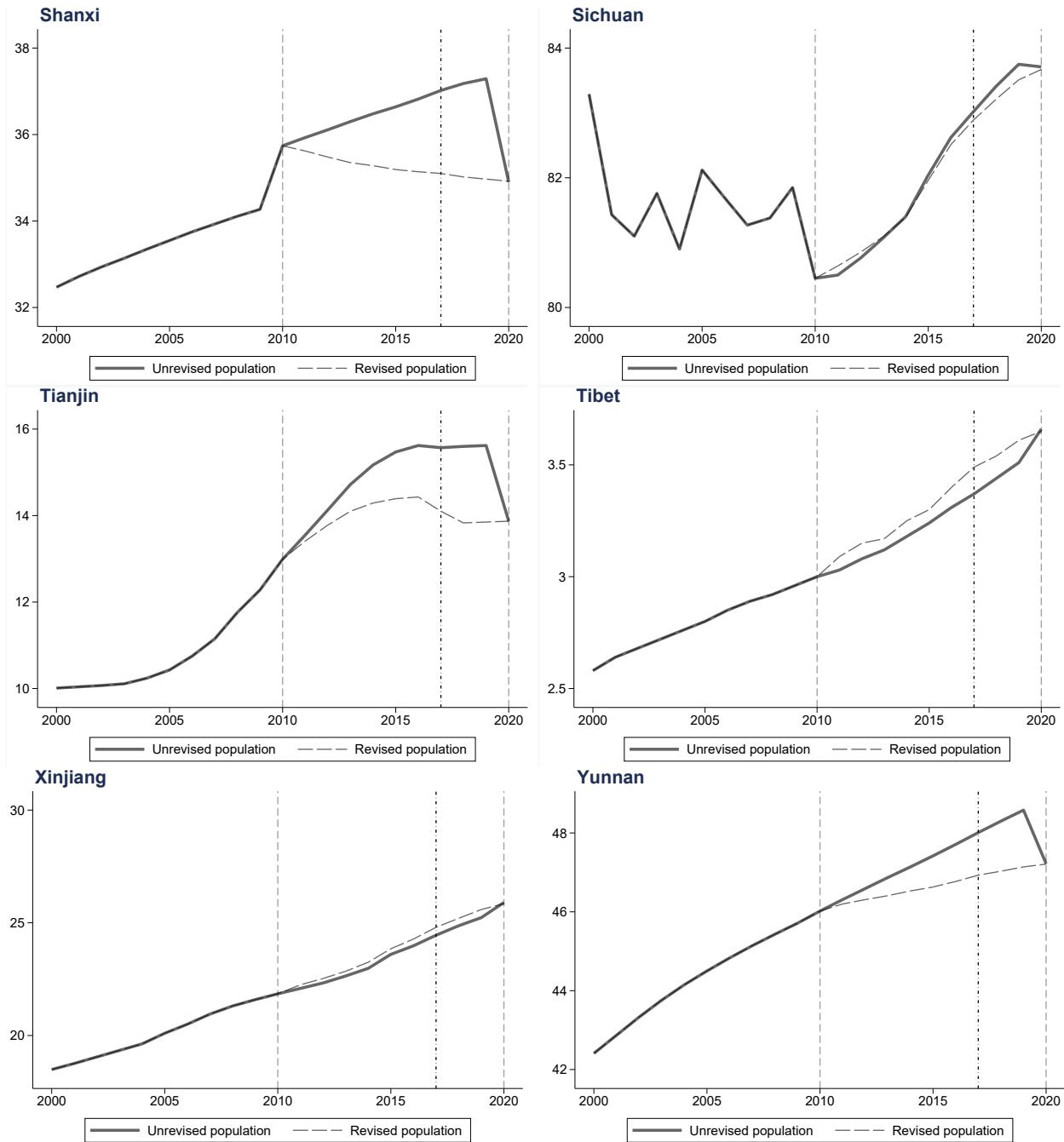


Figure A6. Population path plots for selected provinces [Part 5]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

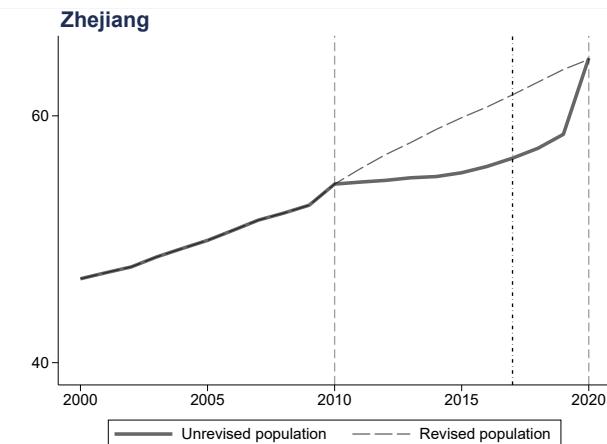


Figure A7. Population path plots for selected provinces [Part 6]. Unrevised and revised population plotted are the raw data from the panel for the period 2000–2020. Dashed vertical lines are for the years 2010 and 2020 which are the end points for the (revised) census population smoothing. Dashed-dotted line is the 2017 GBA announcement.

D. Additional Tables and Figures for Per Capita GDP

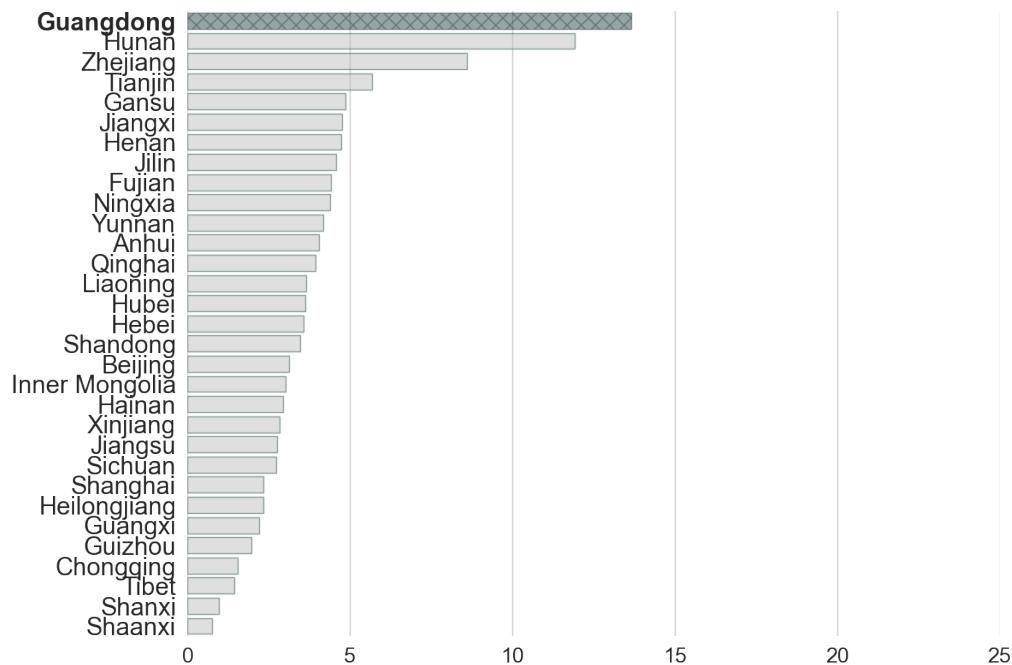


Figure A8. Post-to-pre RMSPE ratio for Per Capita GDP. Untruncated version of [Figure 6](#) for all 31 mainland provinces.

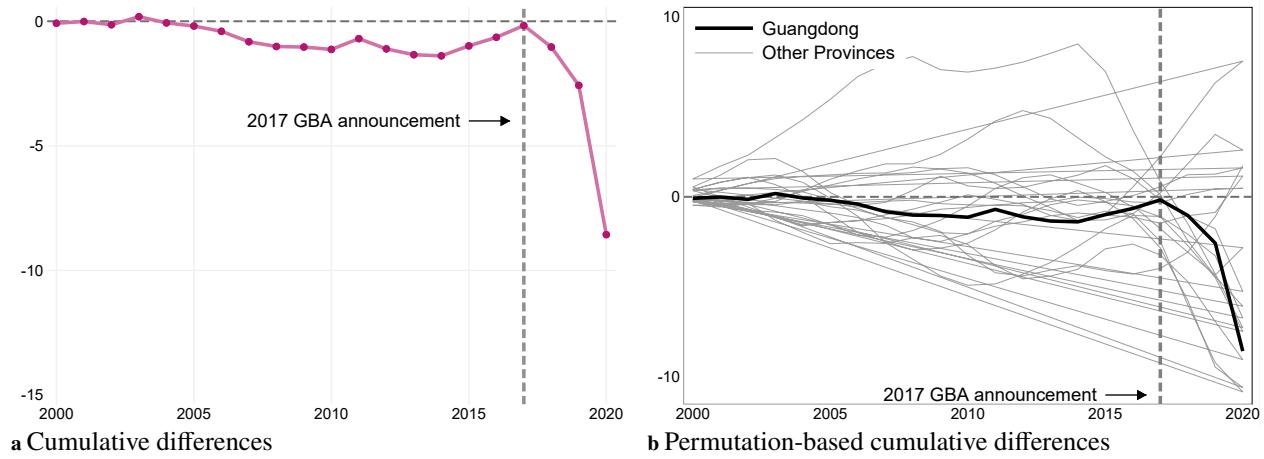


Figure A9. Cumulative differences for Per Capita GDP.

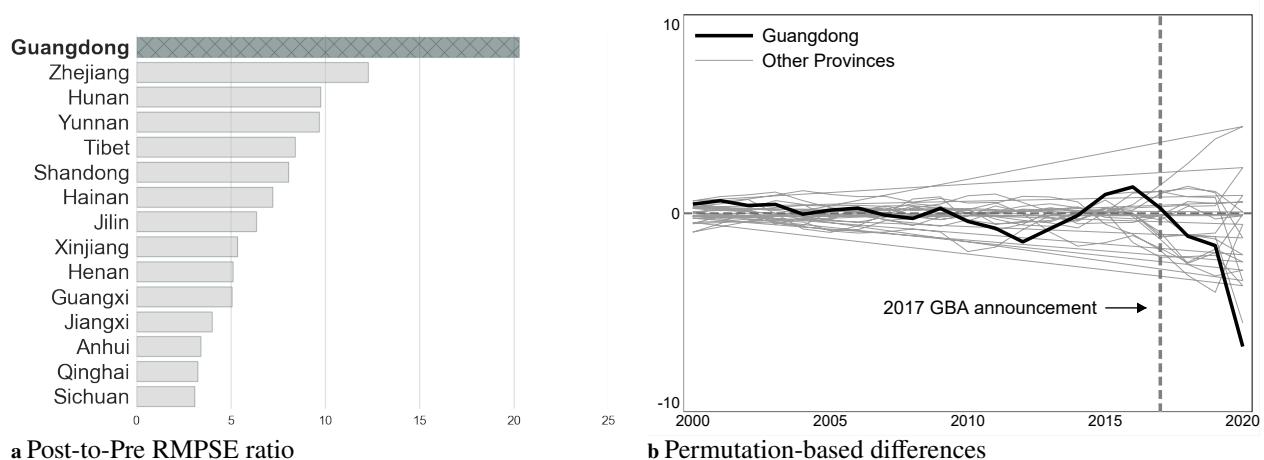
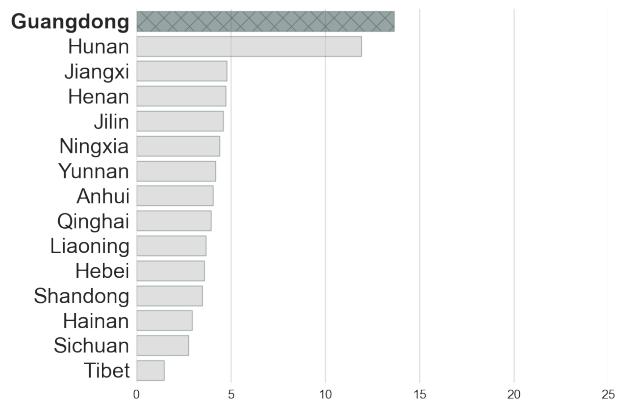
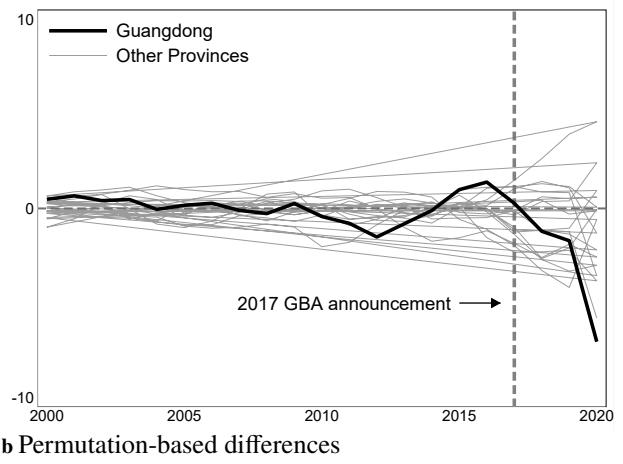


Figure A10. Robustness test for Per Capita GDP with additional predictors: (i) illiteracy rate (percentage of population aged 15 and above), (ii) tertiary institutions (percentage of institutions), (iii) tourist arrivals (normalised by GDP), (iv) realised foreign direct investments (excluding regional investment and normalized by GDP), and (v) realized investments (normalized by GDP). Corresponds to [Figure 6](#).



a Post-to-Pre RMPSE ratio



b Permutation-based differences

Figure A11. Robustness test for Per Capita GDP with a lower constraint violation tolerance in the constrained quadratic optimization. The threshold is set to 0.1^{10} instead of 0.05. Corresponds to [Figure 6](#).

E. Additional Tables and Figures for Population Density

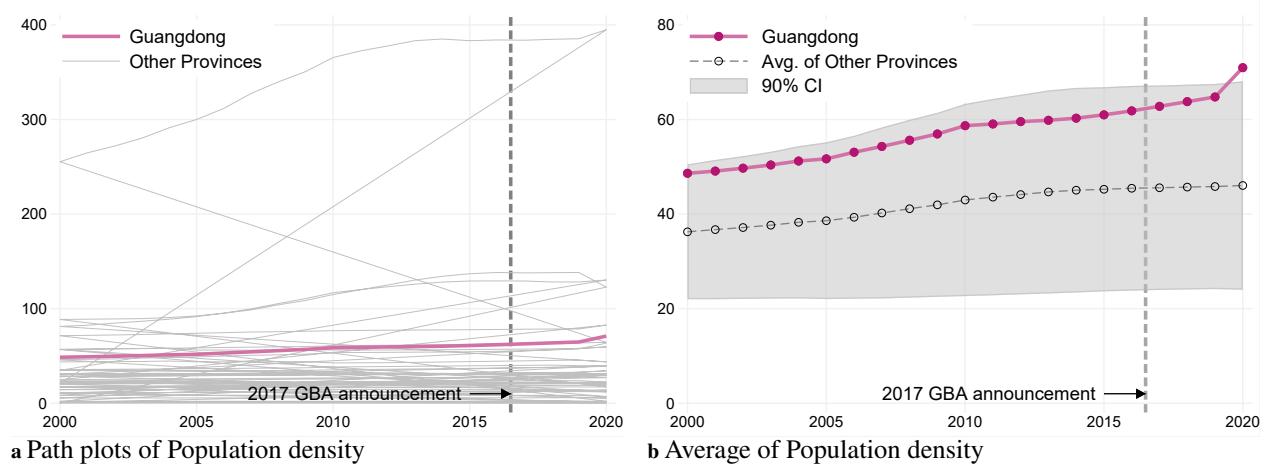


Figure A12. Panel a shows the raw path plots of the province Population density. Gray lines are provinces other than Guangdong. Panel b shows the path plot of Guangdong and the average of the other provinces. Shaded gray area is the 90% confidence interval. Dashed vertical line indicates the 2017 announcement of the GBA.

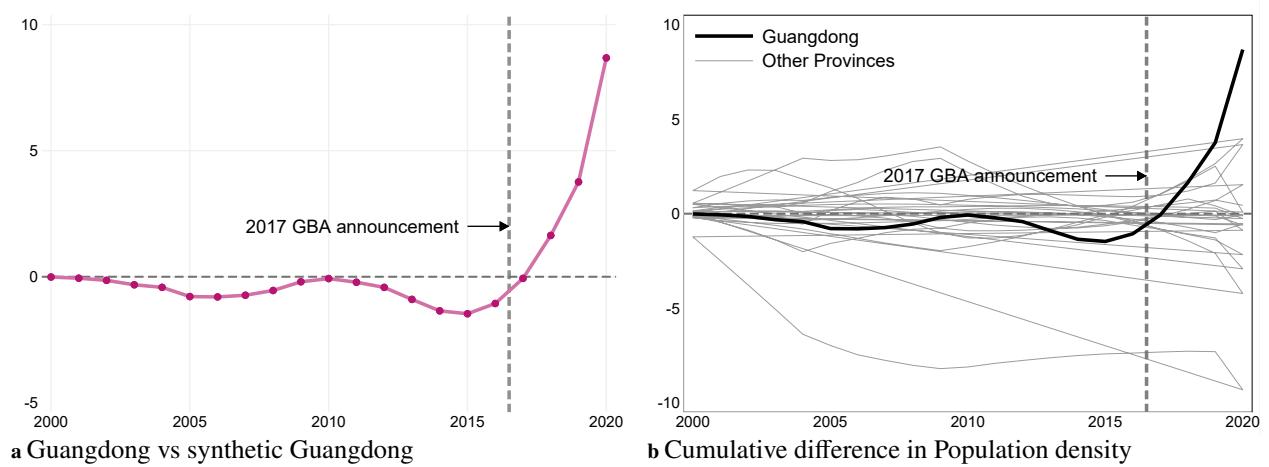


Figure A13. Cumulative differences for Population density.

Table A3. Balance of predictors of Population density before the 2017 GBA announcement.

	Weight	Guangdong	Synthetic Guangdong	Average	Zhejiang (nearest neighbour)	Beijing (next nearest neighbour)
Primary industry	0.001	0.057	0.069	0.127	0.057	0.01
Secondary industry	0.065	0.473	0.472	0.467	0.501	0.231
Tertiary industry	0.006	0.469	0.473	0.456	0.447	0.712
Household income per capita	0.117	20,376	20,388	15,341	24,110	28,704
Visitors	0.006	0.018	0.017	0.015	0.015	0.034
Unemployment rate (urban)	0.001	2.585	3.033	3.537	3.293	1.437
Annual wage	0.803	33,876	33,916	29,353	35,670	57,715

Each row is a predictor used to match the Population density of the 30 mainland provinces to Guangdong. Column 1 is the variable weights. Column 2 is the mean values for Guangdong. Column 3 is the mean predictor values for synthetic Guangdong ($X_P W^*$ from [Equation \(1\)](#)). Columns 4, 5, and 6 are special cases of the synthetic control. Column 4 is where all 30 mainland provinces get equal weights (W is a vector with all elements $1/30$). Column 5 is when the nearest neighbour gets full weight (W all zeroes except with weight 1 for Zhejiang, see [Figure A14](#)). Column 6 is when the second nearest neighbor gets full weight (W all zeroes except with weight 1 for Beijing).

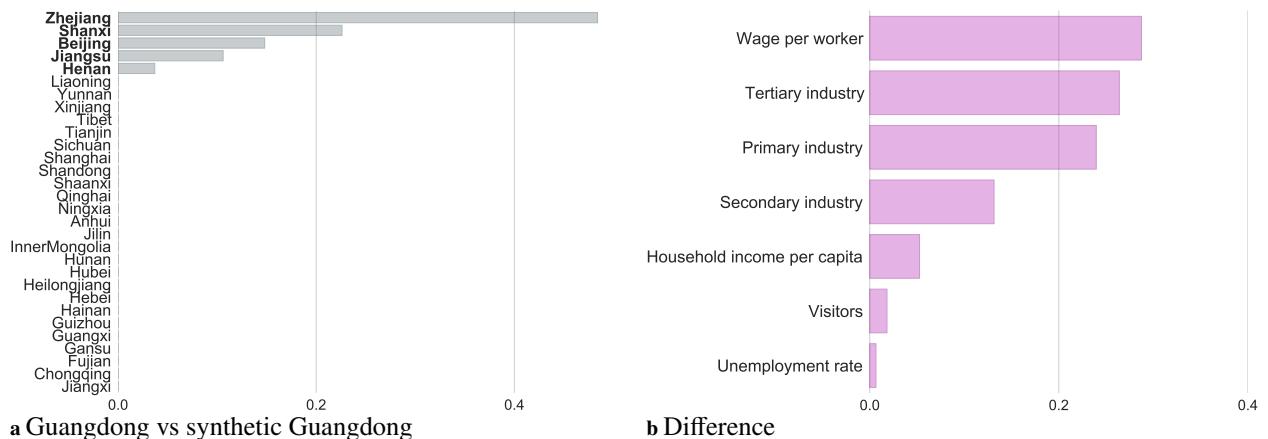


Figure A14. Weights to construct Population density of synthetic Guangdong. Left panel are the synthetic control weights, where weights are W^* from [Equation \(1\)](#) chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Right panel are the predictor variable weights V^* from [Equation \(2\)](#).

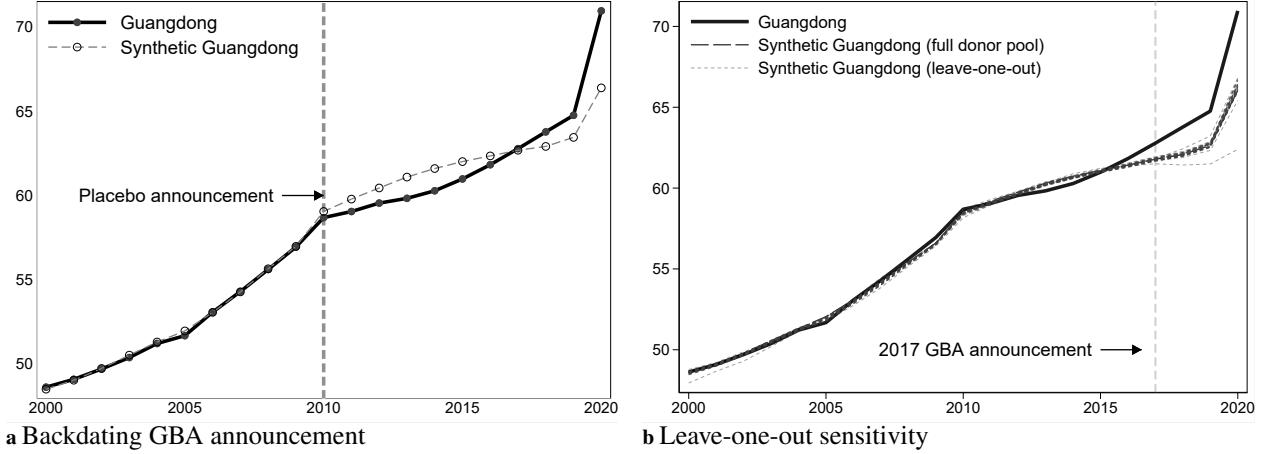


Figure A15. Robustness of synthetic control method for Population density. Left panel repeats the synthetic control method to estimate Population density for synthetic Guangdong but with the GBA announcement backdated to 2010. Right panel shows the results from a leave-one-out synthetic control estimation where in each of 30 iterations one of the donor provinces is taken out of the donor pool before estimation. Dashed vertical line indicates the 2017 announcement of the GBA.

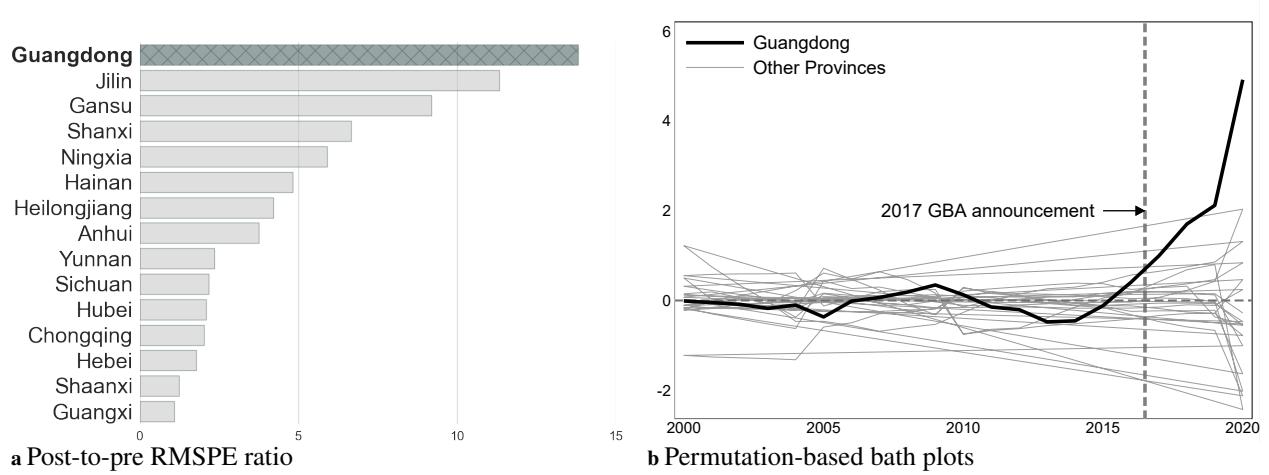


Figure A16. Permutation-based inference for Population density. Left panel shows the distribution of the RMPSE (root mean square prediction error), with the full list in [Table A4](#). Shaded bar is Guangdong. Right panel shows the synthetic control method results by using each of the other 30 mainland provinces in turn as the placebo in gray. The original synthetic control method result for Guangdong from [Figure 8](#) in black. Dashed vertical line indicates the 2017 announcement of the GBA. Both panels retain only mainland provinces with good pre-period fit quality (RMPSE less than the median) for better resolution.

Table A4. Permutation-based inference for population density

Province	Type (1)	Rank (2)	P-value (3)	Pre	Post	Adjustment for fit quality using RMSPE			
				RMSPE (4)	RMSPE (5)	Post-to-Pre RMSPE (6)	Adjusted Rank (7)	Adjusted P-value (8)	Adjusted Z-score (9)
Gansu	Donor	27	0.870968	0.00	0.06	16.43	1	0.03	2.14
Yunnan	Donor	28	0.903226	0.00	0.05	15.13	2	0.06	1.85
Shanxi	Donor	18	0.580645	0.01	0.18	12.53	3	0.10	1.66
Guangdong	Treated	12	0.387097	0.05	0.51	11.08	4	0.13	1.52
Ningxia	Donor	19	0.612903	0.02	0.16	9.75	5	0.16	1.40
Hainan	Donor	17	0.548387	0.02	0.20	8.22	6	0.19	1.30
Jilin	Donor	20	0.645161	0.03	0.14	5.00	7	0.23	1.21
Hebei	Donor	23	0.741935	0.02	0.10	4.61	8	0.26	1.13
Hubei	Donor	25	0.806452	0.03	0.09	2.88	9	0.29	1.06
Anhui	Donor	16	0.516129	0.10	0.27	2.80	10	0.32	0.99
Shaanxi	Donor	26	0.83871	0.03	0.08	2.58	11	0.35	0.93
Liaoning	Donor	15	0.483871	0.16	0.41	2.49	12	0.39	0.86
Tianjin	Donor	6	0.193548	0.57	1.29	2.26	13	0.42	0.81
Beijing	Donor	10	0.322581	0.33	0.72	2.16	14	0.45	0.75
Fujian	Donor	9	0.290323	0.41	0.83	2.05	15	0.48	0.70
Sichuan	Donor	29	0.935484	0.02	0.03	1.99	16	0.52	0.65
Hunan	Donor	24	0.774194	0.05	0.10	1.87	17	0.55	0.60
Jiangsu	Donor	11	0.354839	0.30	0.52	1.71	18	0.58	0.55
Shanghai	Donor	1	0.032258	22.78	36.77	1.61	19	0.61	0.51
Heilongjiang	Donor	14	0.451613	0.28	0.42	1.54	20	0.65	0.46
Henan	Donor	4	0.129032	1.30	1.99	1.53	21	0.68	0.42
Xinjiang	Donor	2	0.064516	1.81	2.72	1.50	22	0.71	0.37
Tibet	Donor	5	0.16129	1.14	1.69	1.48	23	0.74	0.33
Qinghai	Donor	13	0.419355	0.35	0.49	1.42	24	0.77	0.29
Jiangxi	Donor	8	0.258065	0.79	1.12	1.41	25	0.81	0.25
Inner Mongolia	Donor	3	0.096774	1.47	2.04	1.38	26	0.84	0.20
Zhejiang	Donor	7	0.225806	0.85	1.15	1.36	27	0.87	0.16
Guizhou	Donor	22	0.709677	0.08	0.11	1.33	28	0.90	0.12
Shandong	Donor	21	0.677419	0.19	0.13	0.68	29	0.94	0.08
Guangxi	Donor	30	0.967742	0.05	0.03	0.64	30	0.97	0.04
Chongqing	Donor	31	1	0.11	0.02	0.21	31	1.00	0.00

Table shows the synthetic control method by using each of the 31 mainland provinces as the treated unit. Column 1 indicates that Guangdong is the main treated unit of concern. Column 2 and column 3 are the ranking and p-value based on the highest post-period RMSPE (root mean square prediction error). Column 4 and column 5 are the pre and post-period RMSPE. Column 6 is the ratio of the post-to-pre RMSPE which adjusts the post-period RMSPE by the pre-period RMSPEs (see [Figure A16](#) and [Figure A17](#)). Columns 7–9 are the ranking, p-value, and the associated Z-score after the post-period RMSPE is adjusted by fit quality in the pre-period.

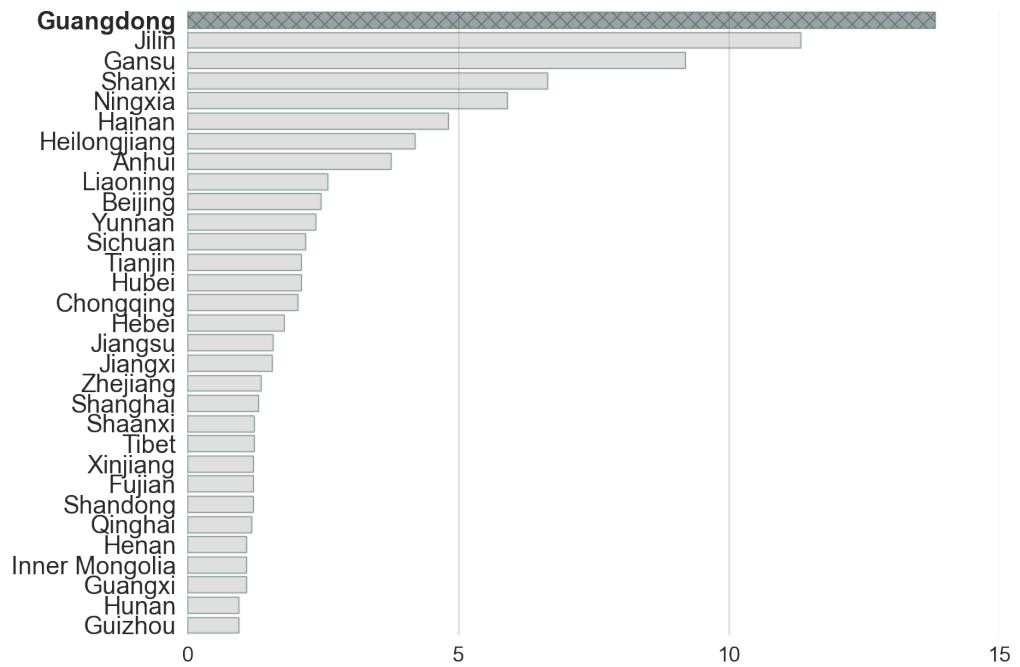


Figure A17. Post-to-pre RMSPE ratio for Population density. Untruncated version of [Figure A16](#) for all 31 mainland provinces.

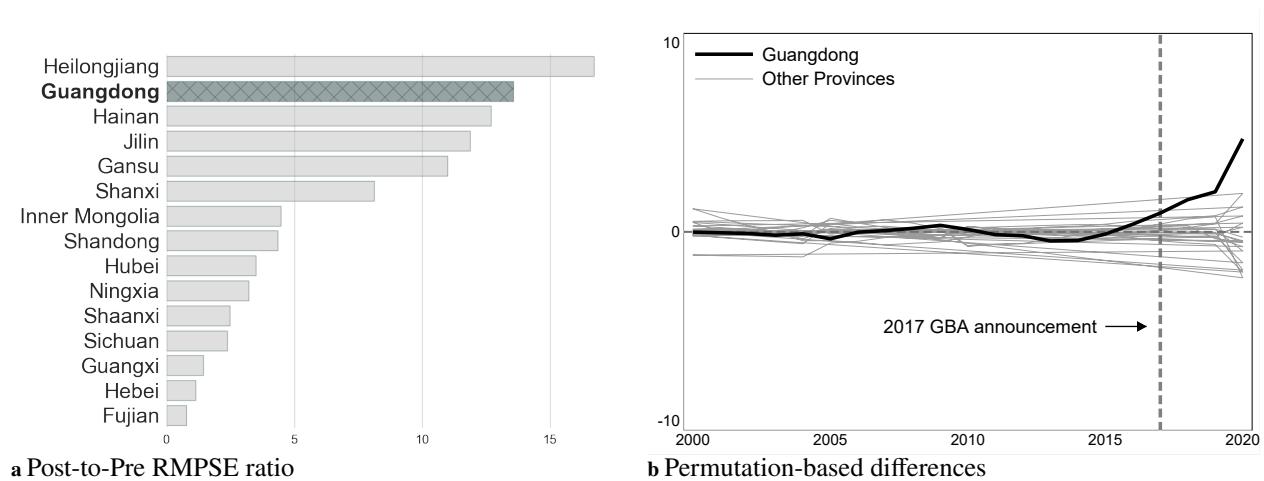
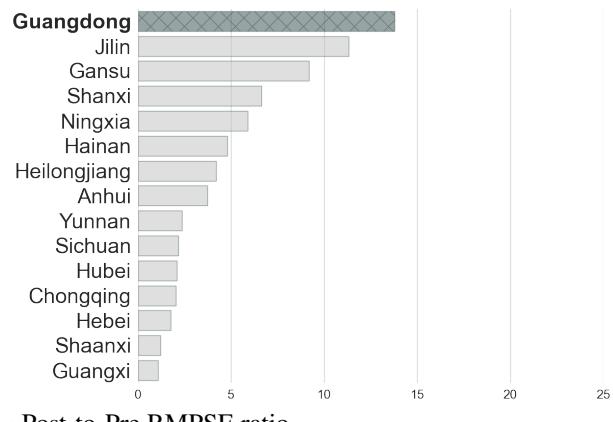
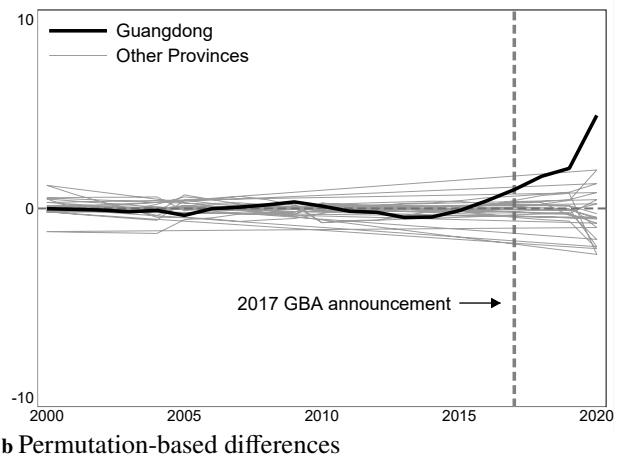


Figure A18. Robustness test for Population density with additional predictors: (i) illiteracy rate (percentage of population aged 15 and above), (ii) tertiary institutions (percentage of institutions), (iii) tourist arrivals (normalised by GDP), (iv) realised foreign direct investments (excluding regional investment and normalized by GDP), and (v) realized investments (normalized by GDP). Corresponds to [Figure A16](#).



a Post-to-Pre RMPSE ratio



b Permutation-based differences

Figure A19. Robustness test for Population density with a lower constraint violation tolerance in the constrained quadratic optimization. The threshold is set to 0.1^{10} instead of 0.05. Corresponds to [Figure A16](#).

F. Additional Tables and Figures for Productivity



Figure A20. Panel a shows the raw path plots of the productivity. Gray lines are provinces other than Guangdong. Panel b shows the path plot of Guangdong and the average of the other provinces. Shaded gray area is the 90% confidence interval. Dashed vertical line indicates the 2017 announcement of the GBA.

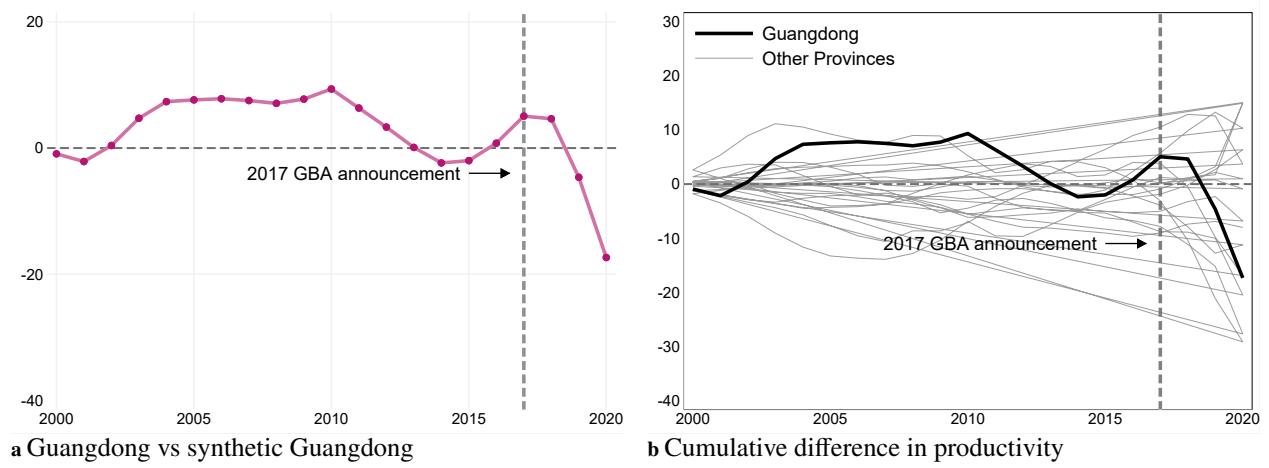


Figure A21. Cumulative differences for productivity.

Table A5. Balance of predictors of productivity before the 2017 GBA announcement.

	Weight	Guangdong	Synthetic Guangdong	Average	Shandong (nearest neighbour)	Beijing (next nearest neighbour)
Primary industry	0.014	0.059	0.071	0.132	0.113	0.011
Secondary industry	0.847	0.479	0.479	0.477	0.588	0.239
Tertiary industry	0.001	0.462	0.516	0.447	0.426	0.699
Household income per capita	0.009	19,214	17,774	14,217	15,500	26,354
Visitors	0.056	0.018	0.017	0.015	0.008	0.036
Unemployment rate (urban)	0.017	2.601	2.775	3.591	3.382	1.447
Annual wage	0.057	30,861	31,900	26,547	24,927	51,925

Each row is a predictor used to match the productivity of the 30 mainland provinces to Guangdong. Column 1 is the variable weights. Column 2 is the mean values for Guangdong. Column 3 is the mean predictor values for synthetic Guangdong ($X_P W^*$ from [Equation \(1\)](#)). Columns 4, 5, and 6 are special cases of the synthetic control. Column 4 is where all 30 mainland provinces get equal weights (W is a vector with all elements $1/30$). Column 5 is when the nearest neighbor gets full weight (W all zeroes except with weight 1 for Zhejiang, see [Figure A22](#)). Column 6 is when the second nearest neighbor gets full weight (W all zeroes except with weight 1 for Beijing).

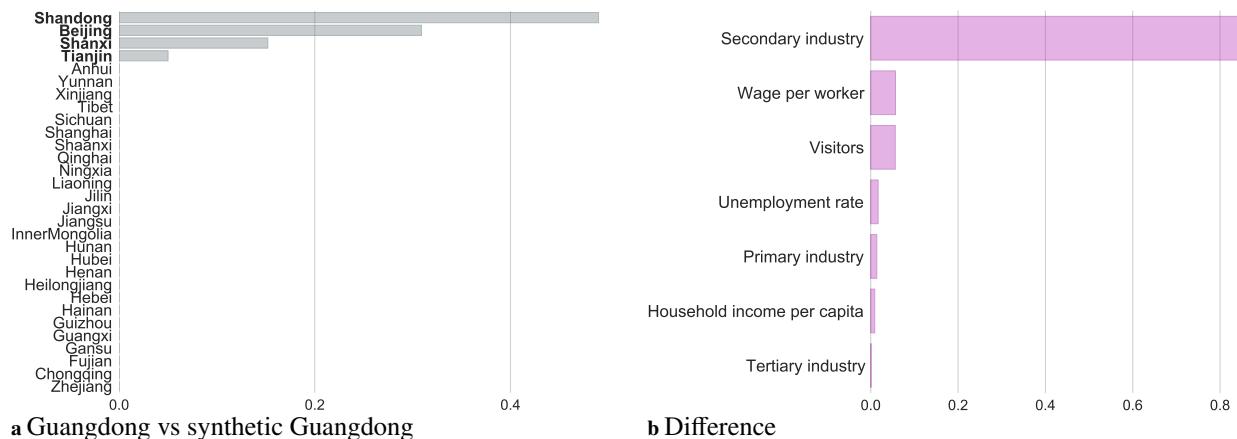
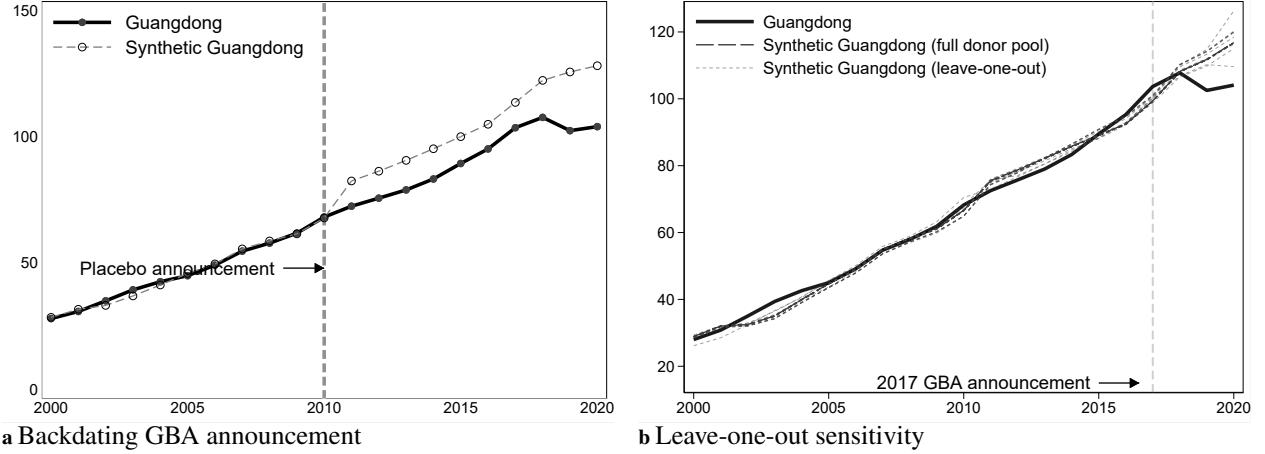


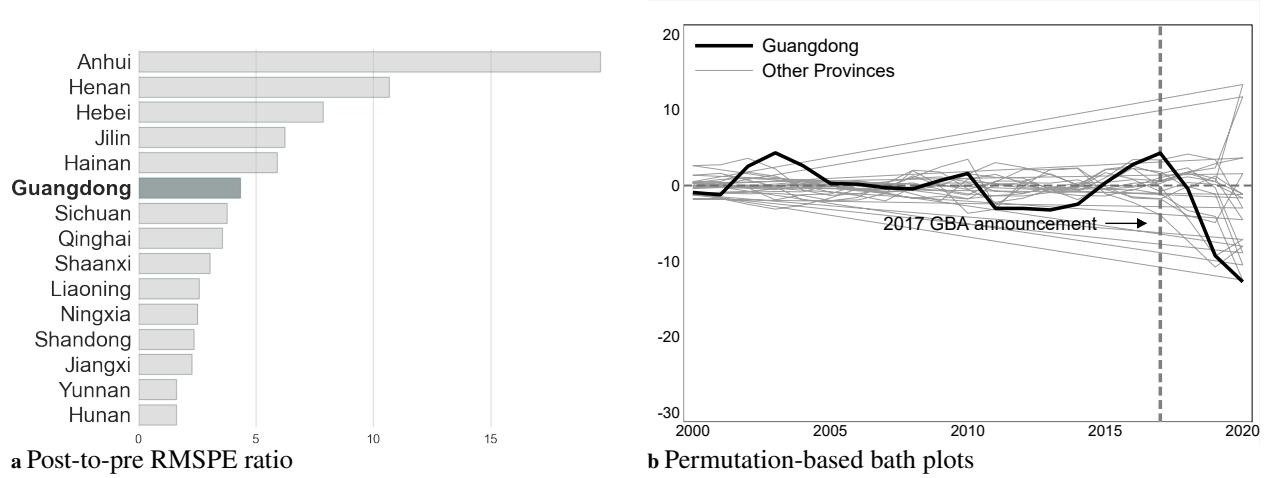
Figure A22. Weights to construct productivity of synthetic Guangdong. Left panel are the synthetic control weights, where weights are W^* from [Equation \(1\)](#) chosen such that synthetic Guangdong best resembles Guangdong in the over the 17 years in the period 2000–2016 before the 2017 GBA announcement. Right panel are the predictor variable weights V^* from [Equation \(2\)](#).



a Backdating GBA announcement

b Leave-one-out sensitivity

Figure A23. Robustness of synthetic control method for productivity. Left panel repeats the synthetic control method to estimate productivity for synthetic Guangdong but with the GBA announcement backdated to 2010. Right panel shows the results from a leave-one-out synthetic control estimation where in each of 30 iterations one of the donor provinces is taken out of the donor pool before estimation. Dashed vertical line indicates the 2017 announcement of the GBA.



a Post-to-pre RMSPE ratio

b Permutation-based bath plots

Figure A24. Permutation-based inference for productivity. Left panel shows the distribution of the RMPSE (root mean square prediction error), with the full list in Table A6. Shaded bar is Guangdong. Right panel shows the synthetic control method results by using each of the other 30 mainland provinces in turn as the placebo in gray. The original synthetic control method result for Guangdong from Figure 8 in black. Dashed vertical line indicates the 2017 announcement of the GBA. Both panels retain only mainland provinces with good pre-period fit quality (RMPSE less than the median) for better resolution.

Table A6. Permutation-based inference for productivity

Province	Type	Rank	P-value	Pre	Post	Adjustment for fit quality using RMSPE			
				RMSPE	RMSPE	Post-to-Pre RMSPE	Adjusted Rank	Adjusted P-value	Adjusted Z-score
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Anhui	Donor	14	0.45	0.39	7.75	19.67	1	0.03	2.14
Henan	Donor	19	0.61	0.66	7.01	10.66	2	0.06	1.85
Hebei	Donor	12	0.39	1.00	7.86	7.86	3	0.10	1.66
Jilin	Donor	10	0.32	1.46	9.11	6.23	4	0.13	1.52
Hainan	Donor	18	0.58	1.22	7.23	5.90	5	0.16	1.40
Inner Mongolia	Donor	7	0.23	2.63	13.84	5.27	6	0.19	1.30
Fujian	Donor	3	0.10	6.67	29.34	4.40	7	0.23	1.21
Guangdong	Treated	9	0.29	2.17	9.43	4.35	8	0.26	1.13
Gansu	Donor	5	0.16	4.29	18.15	4.23	9	0.29	1.06
Hubei	Donor	4	0.13	4.84	19.56	4.04	10	0.32	0.99
Sichuan	Donor	25	0.81	0.76	2.86	3.76	11	0.35	0.93
Qinghai	Donor	22	0.71	1.14	4.05	3.57	12	0.39	0.86
Tibet	Donor	11	0.35	2.72	8.54	3.13	13	0.42	0.81
Shaanxi	Donor	20	0.65	2.04	6.18	3.03	14	0.45	0.75
Guangxi	Donor	8	0.26	3.22	9.62	2.99	15	0.48	0.70
Jiangsu	Donor	1	0.03	17.05	45.00	2.64	16	0.52	0.65
Liaoning	Donor	21	0.68	2.09	5.39	2.58	17	0.55	0.60
Ningxia	Donor	23	0.74	1.55	3.89	2.50	18	0.58	0.55
Shandong	Donor	24	0.77	1.23	2.89	2.35	19	0.61	0.51
Jiangxi	Donor	28	0.90	0.47	1.08	2.26	20	0.65	0.46
Tianjin	Donor	6	0.19	7.99	18.07	2.26	21	0.68	0.42
Beijing	Donor	2	0.06	20.26	43.47	2.15	22	0.71	0.37
Guizhou	Donor	13	0.42	3.81	7.80	2.05	23	0.74	0.33
Heilongjiang	Donor	15	0.48	3.65	7.40	2.03	24	0.77	0.29
Yunnan	Donor	27	0.87	0.77	1.23	1.60	25	0.81	0.25
Hunan	Donor	26	0.84	0.92	1.47	1.60	26	0.84	0.20
Chongqing	Donor	17	0.55	6.14	7.24	1.18	27	0.87	0.16
Xinjiang	Donor	16	0.52	7.06	7.40	1.05	28	0.90	0.12
Shanxi	Donor	29	0.94	2.80	0.39	0.14	29	0.94	0.08
Zhejiang	Donor	30	0.97	6.29	0.35	0.06	30	0.97	0.04
Shanghai	Donor

Table shows the synthetic control method by using each of the 31 mainland provinces as the treated unit. Column 1 indicates that Guangdong is the main treated unit of concern. Column 2 and column 3 are the ranking and p-value based on the highest post-period RMSPE (root mean square prediction error). Column 4 and column 5 are the pre and post-period RMSPEs. Column 6 is the ratio of the post-to-pre RMSPE, which adjusts the post-period RMSPE by the pre-period RMSPEs. Columns 7–9 are the ranking, p-value, and the associated Z-score after the post-period RMSPE is adjusted by fit quality in the pre-period.

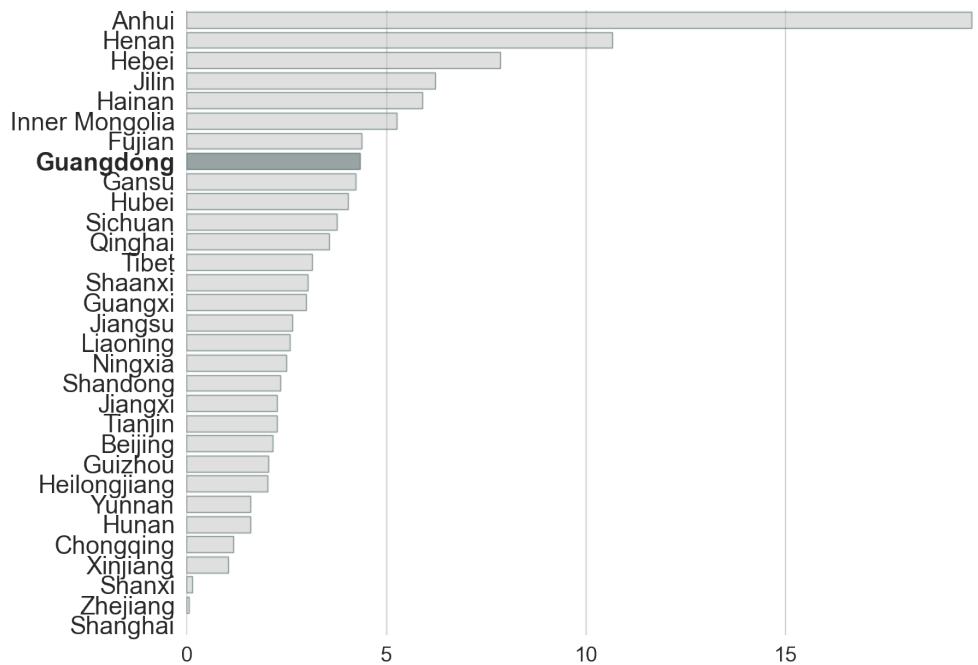


Figure A25. Post-to-pre RMSPE ratio for productivity. Untruncated version of [Figure A24](#) for all 31 mainland provinces.

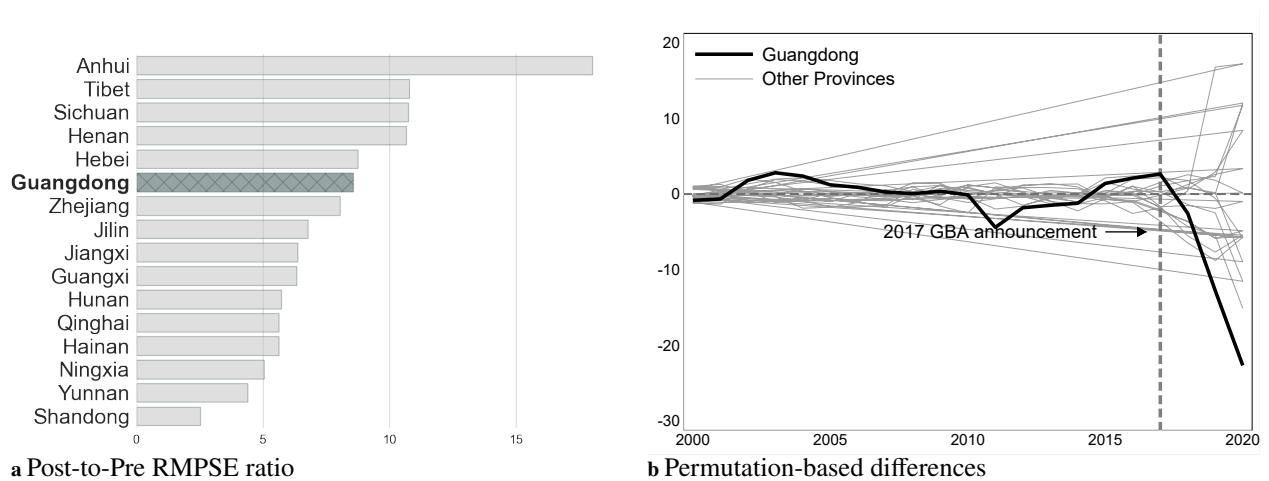
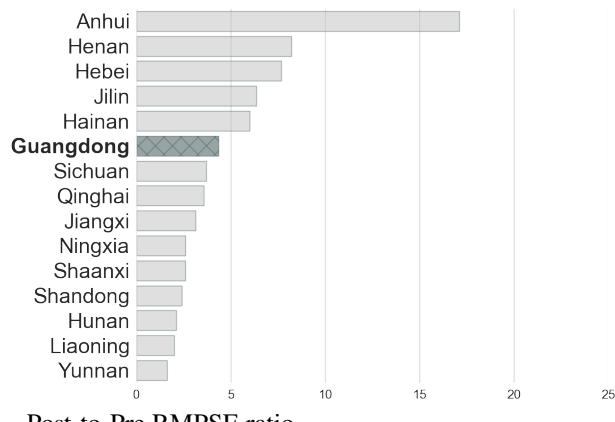
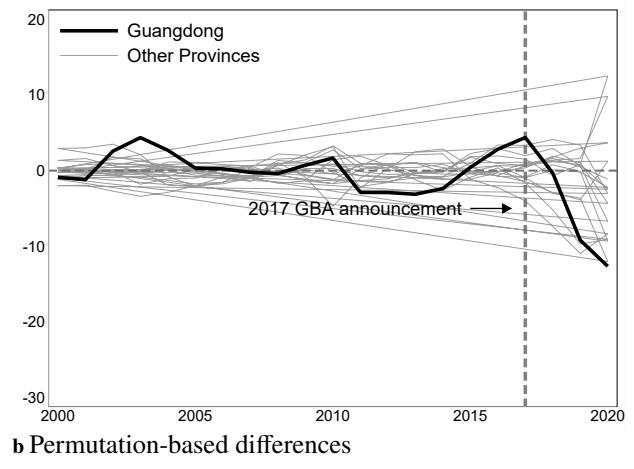


Figure A26. Robustness test for productivity with additional predictors: (i) illiteracy rate (percentage of population aged 15 and above), (ii) tertiary institutions (percentage of institutions), (iii) tourist arrivals (normalized by GDP), (iv) realized foreign direct investments (excluding regional investment and normalized by GDP), and (v) realized investments (normalized by GDP). Corresponds to [Figure A24](#).



a Post-to-Pre RMPSE ratio



b Permutation-based differences

Figure A27. Robustness test for productivity with a lower constraint violation tolerance in the constrained quadratic optimization. The threshold is set to 0.1^{10} instead of 0.05. Corresponds to [Figure A24](#).