What Has New Zealand Gained From The FTA With China?: Two Counterfactual Analyses*

Samuel Verevis† Murat Üngör‡

24 January 2020

Abstract

We investigate the effects of the 2008 New Zealand (NZ)-China free trade agreement (FTA) on exports from NZ to China, and real GDP per capita in NZ using the synthetic control method to estimate the counterfactuals. NZ exports to China were more than 200% higher in 2014 than what they would have had the FTA never been signed. NZ's food and live animals exports to China were more than 180% higher in 2014 than the counterfactual. Both the benchmark and the robustness cases indicate a small but negative effect of the FTA on NZ's real GDP per capita between 2009-2012.

JEL classification: F13, F14, F43, O56.
Keywords: Trade agreements, New Zealand, China, synthetic control method.

*Some parts of this article are based on the first author's master thesis (Verevis, 2018), which was supervised by the second author. The views expressed herein are those of the authors and not necessarily those of the institutes they are affiliated to. We thank an anonymous referee for his/her valuable comments, which greatly helped in improving the paper. The poster of this paper was presented and won the *People's Choice Poster Prize* at the 59th New Zealand Association of Economists Annual Conference in June 2018. An earlier version appeared as the University of Otago Economics Discussion Paper No. 1906.

†Research and Analysis Unit, Ministry of Foreign Affairs & Trade, HSBC Tower 195 Lambton Quay, Wellington 6011, New Zealand. E-mail address: Sam.verevis@mfat.govt.nz

‡Department of Economics, University of Otago, PO Box 56, Dunedin 9054, New Zealand. E-mail address: murat.ungor@otago.ac.nz

1 Introduction

New Zealand (NZ) was the first developed country to sign a free trade agreement (FTA) with China. The NZ-China FTA took effect in October 2008, resulting in a large increase in trade between the two countries. We study this event from the perspective of a policy evaluation problem and investigate the effects of the 2008 NZ-China FTA on trade flows (exports from NZ to China) and income (real GDP per capita in NZ) using the synthetic control method (SCM) to estimate the counterfactuals. We generate several counterfactual scenarios without the FTA to evaluate the performance of our outcome variables, namely, NZ's exports to China and real GDP per capita. Since we have no post-2008 observations on NZ where the FTA with China was not signed, we create a control group by synthesizing the performance of countries similar to NZ. We use a combination of other OECD countries to construct a "synthetic" control NZ, which resembles relevant economic characteristics of the NZ economy before signing the FTA with China. The subsequent economic progress of this "counterfactual" NZ without the FTA with China is compared to the actual experience of the NZ economy.

We ask two questions in this paper. Question one asks: *Have NZ's exports*[1] *to China increased significantly because of the 2008 FTA?* We find that NZ exports to China were more than 200% higher in 2013 and 2014 than what they would have been if NZ had never signed the FTA with China. We also explore trade creation and trade diversion effects of the 2008 FTA for NZ exports. That is, as NZ experienced larger commodity exports going to China, it also experienced declines in exports to other major trading partners, such as Australia and the United States (US). We find that total commodity exports from NZ to the world would have been 22% less than what they were in 2014 had the FTA never occurred. We find that this sector's exports to China would have been 185% less in the counterfactual outcome. The second question we ask is: *How has the FTA with China affected the per capita income of NZ?* We find that NZ's real GDP per capita was 4% lower, on average, during the 2009-2017 period than what they would have been if NZ had never signed the FTA with China.

Our main results, presented in Section 4 and 5, are based on a control group that consists of 24 OECD countries (see Section 4.1). When evaluating our counterfactual results, we report two goodness-of-fit criteria to measure the similarity of our control group to NZ in the pre-treatment period. Our analyses yield good fits (see Figure 3 and Figure 7), indicating that countries in our control group can be considered peer countries of NZ. However, there

[1] All export and import figures in this paper are expressed in nominal US dollars, unless real terms are specified.

are still differences between the countries in our control group and NZ that we don't account for in our benchmark counterfactuals. For example, almost half of the countries in our control group (10 out of 24) belong to the Eurozone monetary union (see Section 6.2). NZ is not a part of this union, therefore, these 10 countries are different from NZ in this aspect. We recognize this as a necessity of robustness, especially for the outcome variable of real GDP per capita. We re-estimate our main results, excluding the Eurozone countries from our control group and find significant (although smaller than our benchmark) effects of the 2008 FTA on NZ's exports to China. This is not surprising considering the significant increase of NZ's export to China in the post-FTA era. In terms of real GDP per capita, the experiment without the Eurozone countries is consistent with our benchmark counterfactual for the 2009-2012 period. However, the experiment without the Eurozone countries suggests gains in real GDP per capita starting with the year 2013. These findings highlight how important the choice of the control groups are in counterfactual exercises and its consequential influence on results.

Our paper, to the best of our knowledge, is one of the first (if not the first) study that provides a systematic quantitative analysis of the 2008 NZ-China FTA on the economic performance of NZ with several counterfactual comparisons.[2] This work is related to a small but growing literature that tries to understand the impact of past trade agreements using the SCM. The two papers most closely related to ours are Hannan (2016) and Hannan (2017). Hannan (2016) explores the impact of past trade agreements using 104 country pairs that had engaged in trade agreements between 1983 and 1995, and finds that substantial gains are generated, with average increases in exports of 80%, and annual growth of 3.8%. Hannan (2017) employs the SCM to determine the impact of trade agreements for 64 Latin American country pairs in the 1989-1996 period. Her results suggest that trade agreements have markedly boosted exports in Latin America, on average by 76.4 percentage points over ten years. Hosny (2012) and Aytuğ et al. (2017) provide more specific examples of trade agreements between countries. Hosny (2012) studies Algeria's trade agreement with the Greater Arab Free Trade Area (GAFTA) and investigates the counterfactual of Algeria signing the FTA in 1998 instead of 2005. Hosny's results suggest that Algeria's trade would have improved in comparison with the counterfactual. Aytuğ et al. (2017) study the effects of EU-Turkey Customs Union (CU) on the Turkish economy. By implementing the SCM the authors find that had Turkey not joined the CU, GDP per capita would have fallen by 13% and exports would have fallen by 38%. By exploring the 'what if' counterfactuals,

[2]There are some descriptive studies on China's impact on the NZ economy. See, for example, Bowman and Conway, 2013a,b; Kendall, 2014. Osborn and Vehbi (2013) provide a quantitative analysis of the impact on NZ of economic growth in China employing a vector autoregressive model.

we aim to dissect the magnitude of bilateral trade agreements on NZ exports and explore the gains from trade, both at the aggregate and sector levels. A further novel contribution to the trade literature is the direct examination of the income effect from trade and its impact on NZ. We do this by studying the evolution of real GDP per capita and estimating the counterfactual real GDP per capita of having never signed the FTA with China. This counterfactual analysis quantifies the effect of the 2008 FTA on NZ's economic well-being, proxied by real GDP per capita that captures the effects of terms of trade changes.

The remainder of the paper is organized as follows. Section 2 presents information regarding how China has become a major trade partner of NZ. Section 3 discusses the SCM as a tool for counterfactual analysis and reviews the related literature. Section 4 presents our quantitative analysis for NZ's exports to China. Section 5 provides the SCM results for real GDP per capita in NZ. Section 6 provides robustness checks and Section 7 concludes.

2 The 2008 NZ-China FTA

For many years, China's development was largely indigenous, mainly because of the country's isolation from the rest of the world. However, over the last three decades China has become an increasingly important part of the global trading system. China has become the leading exporter for merchandise trade and China's accession to the World Trade Organization (WTO) has been marked as an important milestone. China officially started its WTO membership application in 1986 and was formally accepted on 11 December 2001.[3] In accordance with WTO rules, China committed to liberalize further and to better integrate into the global economy. China has signed several FTAs to strengthen international economic cooperation since 2002.[4] NZ was the first developed country to commence FTA negotiations with China. In November 2004, NZ and China launched FTA negotiations and in April 2008, NZ became the first OECD country to successfully conclude FTA negotiations with China.[5] On 7 April 2008, Chinese Premier Wen Jiabao and NZ Prime Minister Helen Clark witnessed the signing of the NZ-China FTA in Beijing, which came into force on 1 October 2008. Since then, NZ has experienced dramatic increases in bilateral trade with China.

Figure 1 shows that China gradually became NZ's main source of (commodity) imports. NZ's imports from China in 1979 were little higher than US$ 37 million, which increased

[3] In 1997 NZ became the first country to agree to China's accession to the WTO by concluding the bilateral negotiations component of that process.

[4] See China FTA Network for details of these FTAs as well as the agreements being negotiated (http://fta.mofcom.gov.cn/english/index.shtml).

[5] Chile-China FTA was signed in November 2005 and entered into force in October 2006. However, Chile was not an OECD country at that time. Chile became an OECD member on 7 May 2010.

to more than US$ 7.7 billion in 2017. NZ's share of imports from China rose from 0.8% in 1979 to around 20% in 2017. In addition, China has become NZ's top commodity export destination in recent years. NZ's export share to China was only 1.9% in 1979, and this share tripled by 2008 at 5.9%. Only five years after that NZ's share of exports to China tripled again, to 20.7% in 2013. Exports to China increased from around US$ 90 million in 1979 to almost US$ 8.5 billion in 2017.

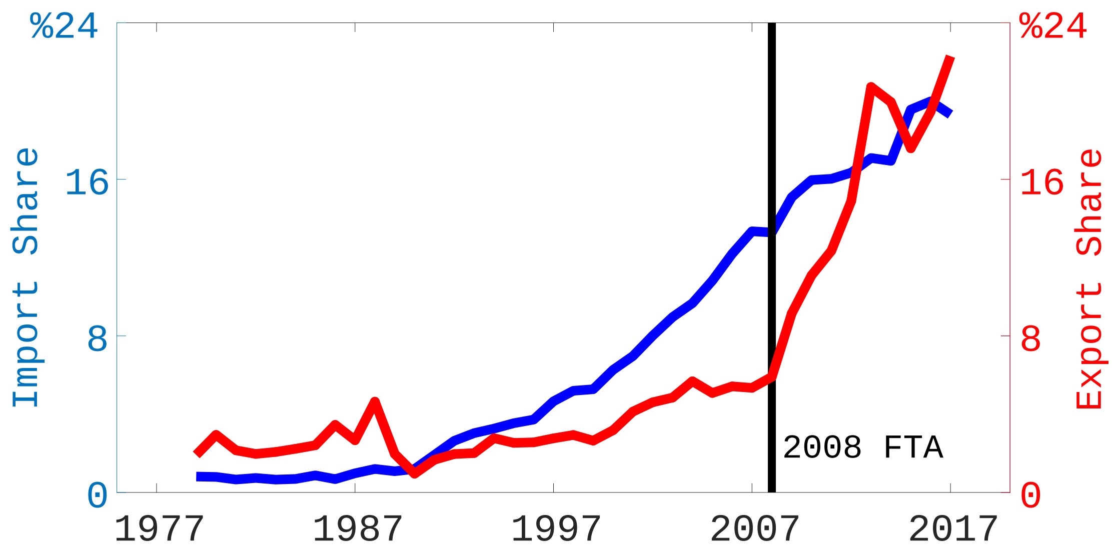

Figure 1: NZ's exports to and imports from China, 1979-2017

Source: `https://comtrade.un.org/data`

Figure 2 demonstrates China's increasing importance for NZ as a major trading partner for total merchandise goods.[6] Panel (a) in Figure 2 shows NZ's three most important partners in terms of total commodity imports. In 1981, Australia and the US together supplied 36.2% of NZ's commodity imports. That share declined to 22.9% in 2017 as China increased its share of NZ's imports from 0.7% in 1981 to almost 20% in 2017. China gradually became NZ's main source of imports. Panel (b) in Figure 2 plots NZ's three most important trade partners in terms of total commodity exports. In 1981, NZ's share of exports to Australia and the US totalled 26.5%, whereas the corresponding share of exports to China was only 2.2%. The share of exports going to China had gradually increased to 5.3% in 2007. After the FTA was signed in 2008, there was a marked increase in NZ's export share going to China. However, the share of exports going to the US declined from 11.5% in 2007 to 9.9% in 2017. The corresponding figures for Australia were 21.9% in 2007 and 16.4% in 2017.[7]

[6] In terms of goods and services (not just commodities), Australia remains NZ's top trading partner, with China NZ's second biggest trading partner (https://www.mfat.govt.nz).

[7] However, NZ remains a relatively unimportant trading partner for China. This is not surprising con-

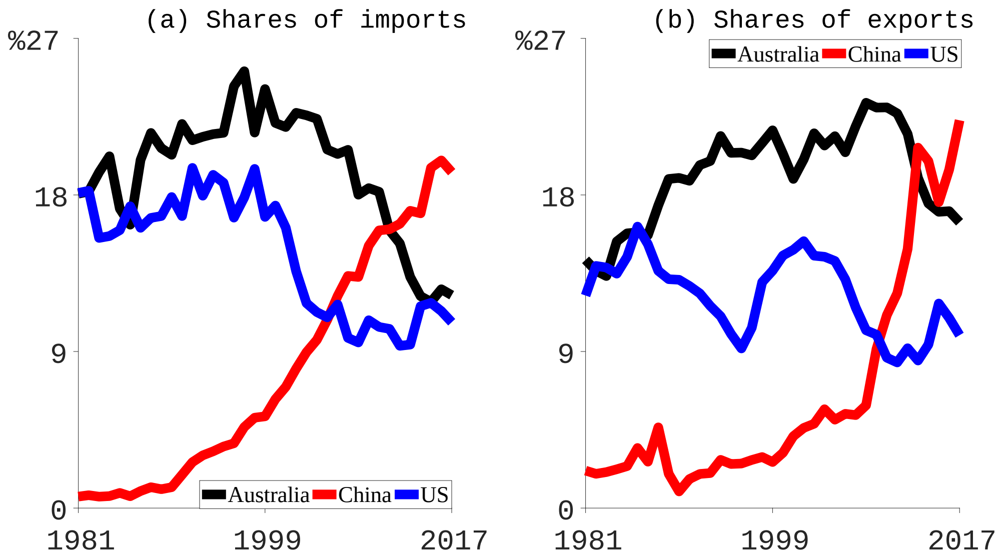

Figure 2: Shares of NZ total commodity trade (%), 1981-2017

Source: `https://comtrade.un.org/data`

NZ exporters have benefitted from the gradual removal of tariffs on most NZ's exports to China. In 2005 China's average applied tariff was around 10% across all products, with a higher average of 14.6% applied to agriculture product (MFAT, 2007). The FTA resulted in the removal of these tariffs and reduction of other impediments to bilateral trade over time. Tariffs are now eliminated for over 97% of NZ goods exports to China. In 2018, all exports other than dairy (some products remain subject to tariffs and safeguards that will be phased out by 2024), and a small number of products that were excluded from the FTA are eligible for tariff-free access into China. In 2018, all imports from China are eligible for tariff-free access. NZ and China agreed in November 2016 to launch negotiations to upgrade the NZ-China FTA. Negotiations commenced in April 2017. Nine rounds of negotiations were held in both NZ and Beijing. NZ and China announced the conclusion of the negotiations to upgrade the FTA between them on 4 November 2019. The upgraded FTA will add new provisions to the existing FTA.[8]

sidering the relative sizes of each economy. NZ's population is less than 5 million people, whereas China's population is more than 1.3 billion people. China was a US$ 12.2 trillion economy in 2017 making it the second largest in the world in terms of nominal GDP. On the other hand, NZ's GDP was little higher than US$ 0.2 trillion in 2017 (World Bank, 2018).

[8] `https://www.mfat.govt.nz/en/trade/free-trade-agreements`

3 Empirical Strategy

3.1 The Potential Outcome Model of Causal Inference

Program evaluation methodologies have long been used to measure the effect of different economic or political interventions (treatments). The central problem in causal inference is evaluating how exposure of a treatment on a set of units (or individuals) has affected their outcomes.[9] Let us illustrate this using the potential outcome set up (Rubin, 1974; Holland, 1986; Imbens and Rubin, 2015). Consider $J+1$ units over $t = 1, ..., T_0, T_0+1, ..., T$ periods. The first unit is affected uninterruptedly by an intervention in period T_0+1 until period T, after an initial pre-intervention period $t = 1, ..., T_0$. The leftover J units are the controls that form the so-called 'donor pool', and they are not affected by the intervention (Gardeazabal and Vega-Bayo, 2017). Let $Y_{j,t}$ denote the outcome variable of unit j at period t. $Y^1_{j,t}$ and $Y^0_{j,t}$ denote the potential outcome of unit j at time t under treatment and in the absence of treatment, respectively. We can write $Y_{j,t} = D_{j,t}Y^1_{j,t} + D_{j,t}Y^0_{j,t}$, where $D_{j,t}$ is a dummy variable that takes value 1 if unit j is under treatment at time t and value 0 otherwise. The identification problem is that the treatment effect depends on the potential outcome in both states ($D_{j,t} = 0$ and $D_{j,t} = 1$), while we can only observe realizations under one of the potential treatment statuses.

3.2 The Synthetic Control Approach

Building on an idea in Abadie and Gardeazabal (2003), Abadie et al. (2010) develop the synthetic control method (SCM) that exploits variation over time in the outcomes of units that are either exposed to treatment only after some period or that are never exposed (see also Abadie et al., 2015). A primary reason to use this method is to control for the effect of unobservable factors that have an impact on the common time trend in the treatment and control groups (Abadie et al., 2010; Acemoglu et al., 2016). Since its introduction, the SCM has seen a range of applications in economics, political science, and international relations.[10]

[9] See Imbens and Wooldridge (2009) and Athey and Imbens (2017) for comprehensive reviews of the recent developments in program evaluation literature.

[10] Examples of such applications include openness and trade liberalization policies (Nannicini and Billmeier, 2011; Billmeier and Nannicini, 2013; Ritzel and Kohler, 2017), impact of trade agreements (Hosny, 2012; Hannan, 2016, 2017; Aytuğ et al., 2017), impact of joining a currency union (Saia, 2017; Puzzello and Gomis-Porqueras, 2018), economic regimes/political stability (Grier and Maynard, 2016; Meyersson, 2017; Jales et al., 2018; Matta et al., 2019), natural disasters (Coffman and Noy, 2012; Cavallo et al., 2013; Mideksa, 2013; Barone and Mocetti, 2014; Mohan, 2017), terrorism, civil wars, crime, and political risks (Montalvo, 2011; Pinotti, 2015; Singhal and Nilakatan, 2016; Bilgel and Karahasan, 2017; Bove et al., 2017; Costalli et al., 2017; Bove and Elia, 2018), health economics (Bilgel and Galle, 2015; Krief et al., 2016), economic sanctions (Gharehgozli, 2017), migration (Powell et al., 2017), and natural resource discoveries (Smith, 2015).

In the words of Athey and Imbens (2017, p. 9), the synthetic control approach is "arguably the most important innovation in the policy evaluation literature in the last 15 years."

Suppose that we observe data for $(J+1) \in \mathbb{N}$ countries. We also assume that there is a treatment that affects only country 1 from period T_0+1 to period T uninterruptedly, where $T_0 \in (1, T) \cap \mathbb{N}$. In other words, without loss of generality, we assume that first country signed a FTA with China, so that we have J remaining countries that serve as potential controls. Let the scalar $Y^0_{j,t}$ be the potential outcome that would be observed for country j in period t if there were no treatment for $j \in 1, ..., J+1$ and $t \in 1, ..., T$. Let the scalar $Y^1_{j,t}$ be the potential outcome that would be observed for country j in period t if country j received the treatment (i.e., signed the FTA with China) from period T_0+1 to period T. We define $\alpha_{j,t} = Y^1_{j,t} - Y^0_{j,t}$ as the treatment effect for country j in period t and $D_{j,t}$ as the treatment indicator that assumes value 1 if country j is treated in period t and value 0 otherwise. The observed outcome for country j in period t is given by

$$Y_{j,t} = Y^0_{j,t}(1 - D_{j,t}) + Y^1_{j,t} D_{j,t}. \tag{1}$$

Because only the first country is exposed to the intervention and only from period T_0+1 to T, we have that

$$D_{j,t} = \begin{cases} 1, & \text{if } j = 1 \text{ and } t > T_0 \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The goal is to estimate the treatment effect on the treated, i.e., $\alpha_{1,t} = Y^1_{1,t} - Y^0_{1,t}$. Recall that we have defined $\alpha_{j,t} = Y^1_{j,t} - Y^0_{j,t}$. This guarantees that we only need to estimate the counterfactual $Y^0_{1,t}$ to identify $(\alpha_{1,T_0+1}, ..., \alpha_{1,T})$ because $Y^1_{1,t}$ is observable for $t > T_0$ (Ferman et al., 2018). Since $Y^0_{1,t}$ for $t > T_0$ is not observed, the main idea of the SCM is to consider a weighted average of the control countries to construct a proxy for this counterfactual.

Let $\mathbf{W} = (w_2, ..., w_{J+1})'$ be a collection of weights, with $w_j \geq 0$ and for $j = 2, ..., J+1$ and $w_2 + w_3 + ... + w_{J+1} = 1$. Each value of $\mathbf{W}$ represents a potential synthetic control. Let $\mathbf{X_1}$ be a $(K \times 1)$ vector of pre-intervention (pre-FTA) values of K predictors for the treated country (NZ). Let $\mathbf{X_0}$ be a $(K \times J)$ matrix which contains the values of the same variables for the J possible control countries. Let $\mathbf{V}$ be a diagonal matrix with nonnegative components. The values of the diagonal elements of $\mathbf{V}$ reflect the relative importance of the different predictors (Abadie and Gardeazabal, 2003). The vector of weights $\mathbf{W}^* = (w^*_2, ..., w^*_{J+1})'$ is chosen to minimize $(\mathbf{X_1} - \mathbf{X_0 W})'\mathbf{V}(\mathbf{X_1} - \mathbf{X_0 W})$ subject to $w^*_j \geq 0$ for $j = 2, ..., J+1$ and $\sum_{j=2}^{J+1} w^* = 1$. The synthetic control estimator of the effect of the treatment for the treated country in post-intervention period t $(t > T_0)$ is $\hat{\alpha}_{1,t} = Y_{1,t} - \sum_{j=2}^{J+1} w^*_j Y_{j,t}$ (Abadie and

Cattaneo, 2018). Arguably, the choice of $\mathbf{V}$ could be subjective and Abadie and Gardeazabal (2003) and Abadie et al. (2010, 2015) propose data-driven selectors of $\mathbf{V}$.

Abadie et al. (2010, 2015) use a minimum distance approach, combined with the restriction that the resulting weights are nonnegative and sum to one.[11] The synthetic control is calculated as a convex combination of the countries in the donor pool and best replicates the outcome variable of the treated country in the pre-intervention period. As long as the weights reflect structural parameters that would not vary in the absence of the 2008 China-NZ FTA, the synthetic control provides a counterfactual scenario for the evolution of the NZ economy in the absence of this FTA (Pinotti, 2015).[12]

3.2.1 Placebo Tests

In order to ensure that a particular synthetic control estimate reflects the impact of the 2008 FTA, we perform a series of in-space placebo tests by applying the SCM on countries in the donor pool that did not sign a FTA with China. These tests compare the magnitude of the estimated effect on the treated country with the size of those obtained by assigning the treatment randomly to any (untreated) country of the donor pool (Olper et al., 2018). This allows us to assess whether the effect estimated by the synthetic control for the country affected by the FTA is large relative to the effect estimated for a country chosen at random (which was not exposed to the 2008 FTA). Large sample asymptotic inference tests are not appropriate in our case because our data are of small sample size and identification of the treatment effect arises from the change in policy by a small group of countries. However, placebo tests à la Abadie and Gardeazabal (2003), Bertrand et al. (2004), Abadie et al. (2010, 2015) can be used to evaluate the significance of treatment effects (Adhikari and Alm, 2016).

Abadie et al. (2015) propose an inference procedure where they permute which unit is assumed to be treated and estimate, for each $j \in \{2, ..., J+1\}$ and $t \in \{1, ..., T\}$, $\hat{\alpha}_{j,t}$ as described above. Then, they compute the test statistic for the ratio of the mean squared

[11] Abadie and Cattaneo (2018) note that in general (that is, ruling out certain degenerate cases) if $\mathbf{X_1}$ does not belong to the convex hull of the columns of $\mathbf{X_0}$, then $\mathbf{W}^*$ is unique and sparse, meaning that $\mathbf{W}^*$ has only a few nonzero elements. Doudchenko and Imbens (2016) propose a more general class of synthetic control estimators, which allow the weights to be negative and do not necessarily restrict the sum of the weights. Abadie and L'Hour (2019) address the problem of multiplicity of solutions, which is finding a synthetic control that best reproduces the characteristics of a treated unit may not have a unique solution. A recent research agenda focuses on (robust) generalization of the SCM (see, for example, Xu, 2017; Amjad et al., 2018; Becker and Klößner, 2018).

[12] The SCM builds on difference-in-differences (DID) estimation, but when constructing the counterfactual, SCM puts more weight on donor countries that closely resemble the treated country, whereas the DID approach assigns equal weight to each donor country. Therefore, the SCM uses systematically more attractive comparisons than those of the traditional DID estimation (Athey and Imbens, 2017).

prediction errors (i.e., the ratio of post/pre treatment mean squared prediction errors). They name this ratio as RMSPE. In addition, they propose to calculate a p-value:

$$p = \frac{\sum_{j=1}^{J+1} \mathbb{1}[RMSPE_j \geq RMSPE_1]}{J+1}, \tag{3}$$

where $\mathbb{1}[\oplus]$ is the indicator function of event $[\oplus]$, and reject the null hypothesis of no effect if p is less than some pre-specified significance level (Ferman et al., 2018).[13]

3.2.2 Pre-treatment Fit Index

We also use the pre-treatment fit index to assess whether the comparison country created by the SCM is a 'good' counterfactual following the work of Adhikari and Alm (2016) and Adhikari et al. (2018). We use this index to assess the overall quality of the pre-treatment fit. The advantage of the pre-treatment fix index is that it normalizes the RMSPE. This makes it possible to compare the fit between different SCM models. This is done by defining the benchmark RMSPE, which is obtained from the zero-fit model.[14] The fit index is defined as the ratio of the RMSPE to the benchmark RMSPE:

$$Fit\,Index = \frac{RMSPE}{Benchmark\,RMSPE}. \tag{4}$$

The range for this index is [0,U], where U is the finite upper bound, meaning that the RMSPE is equivalent to the RMSPE obtained when the difference between the treated and the synthetic unit is U percent for each pre-treatment year (Adhikari et al., 2018). If the index is 0, then the fit is perfect. An index greater than 1 indicates a poor fit and synthetic units should be discarded. We report the pre-treatment fit index and the RMSPE values. If the pre-intervention 'fit' is good, the post intervention counterfactual is likely to be more accurate.

[13] Although the p-value from this placebo test lacks a clear statistical interpretation, this test is commonly used in the SCM applications. Ferman and Pinto (2017), Hahn and Shi (2017), Ferman et al. (2018), Firpo and Possebom (2018) provide comprehensive discussions using both theoretical and Monte Carlo methods.

[14] $RMSPE = \sqrt{\frac{1}{T_0}\sum_1^{T_0}\left(Y_{1,t} - \sum_{j=2}^{J+1} w_j^* Y_{j,t}\right)^2}$. $Benchmark\,RMSPE = \sqrt{\frac{1}{T_0}\sum_1^{T_0}\left(Y_{1,t}\right)^2}$.

4 A Counterfactual Design for NZ Exports

4.1 Construction of the Synthetic NZ

We use annual country-level data for the 1991-2015 period. This gives us a pre-intervention period of 17 years (1991-2007). We construct a synthetic NZ, made from a donor pool of 24 OECD member countries: Australia, Austria, Canada, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, the United Kingdom (UK), and the US. We start with the 34 non-NZ OECD countries; however, we automatically exclude Chile since Chile had an FTA with China before 2008. We exclude Turkey and Mexico because of their middle-income status and these two countries cannot be named as peer countries to the NZ economy. We also exclude the post-communist countries such as Czechia, Estonia, Latvia, Slovakia, and Slovenia. Belgium and Luxembourg are excluded because of their small sizes and limited trade flow observations.

The outcome variable is (nominal) exports to China, measured in current US dollars. Motivated by the gravity model[15] as predictors of trade flows, we use nominal GDP, GDP per capita, population and bilateral exchange rates, and population weighted distance between country i and China (see Appendix A.1). Data for trade flows are from the World Integrated Trade Solution (WITS) database, which uses the United Nations' Comtrade database. WITS reports several nomenclatures and we use the Standard International Trade Classification (SITC2-Division) data to extend the sample period. We also source all of our SITC2 commodity level export values from the WITS, along with total exports to the world. All are measured in current US dollars. From the World Bank's World Development Indicators (WDI), we retain the variables for GDP (in current US dollars) and population (World Bank, 2018). Exchange rates are from the OECD.[16] Exchange rates are measured in national currency per US dollar. This is used to construct the variable national currency per Chinese yuan, which is the national currency reflected in Chinese yuan on an annual level. Finally, we use the population weighted distance from the GeoDist dataset of the CEPII.[17]

[15]The gravity equation has been used as a workhorse for analysing the determinants of bilateral trade flows since Tinbergen (1962). Head and Mayer (2014) provide a historical review of the fundamental equation and discuss theory-consistent estimation, covering number of alternative techniques. Estimating the gravity equation is beyond the scope of this current paper.

[16]`http://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE4`

[17]The CEPII (The Centre d'Études Prospectives et d'Informations Internationales) gathers and harmonizes data from different sources, produces indicators and statistical measures (`http://www.cepii.fr/cepii/en/bdd_modele/bdd.asp`). We take the related CEPII variables from their dataset GeoDist (`http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=6`). GeoDist provides several geographical variables, bilateral distances measured using city-level data to account for the geographic distribution of population inside each nation. We use the population weighted distance.

The transparency of the SCM allows us to observe which donor countries were chosen and the particular weights applied for the construction of the synthetic NZ. Five countries receive positive weights: Canada (4.8%), Iceland (13.9%), Japan (0.7%), Portugal (78.5%) and Sweden (2.2%). All other countries obtain zero weights within the donor pool.[18] Table 1 displays the pre-treatment characteristics of the NZ economy and compares them with the characteristics of the sample average and the characteristics of the weighted average of the 24 OECD countries using the synthetic control weights. We include three lagged values of exports in 1995, 1999, and 2006 as additional predictors to minimize the possible problem of time-constant restriction on predictors (Chung et al., 2016). These lagged values can also explain the time trend of exports to China. Overall, the synthetic NZ provides a much better comparison than using a sample average of the OECD countries in the donor pool.

Table 1: Trade predictor means before the NZ-China FTA

	NZ	Synthetic NZ	OECD Sample
GDP	7.24E+10	1.88E+11	1.10E+12
GDP per capita	18517	17865	32233
Population	3847029	10600000	35119669
Bilateral exchange rate	0.23	1.52	8.38
Weighted distance	10241	9707	7876
Lagged exports 1995	0.35	0.33	2.76
Lagged exports 1999	0.33	0.31	3.08
Lagged exports 2006	1.21	1.24	13.2
RMSPE		0.193	

Note: All variables are averaged over 1991-2007.

4.2 Counterfactual and Inference Tests

Figure 3 displays the time paths of NZ's exports to China and its synthetic counterpart. During the pre-intervention period the synthetic NZ closely reproduces the exports of NZ to China. The pre-treatment fit index is 0.037. The estimated effect of the FTA on NZ's exports is given by the difference between actual NZ exports and that of its synthetic counterpart, shown in Figure 3. In the post-intervention between 2009-2015, NZ's actual exports to China were more than 120%, on average, higher than the synthetic counterparts. Both the immediate divergence and magnitude of the estimated effect, relative to the counterfactual indicates how important and impactful the FTA was on NZ's exports to China.

[18]Like for nearest neighbor matching estimators, most of the synthetic control weights are equal to zero and a small number of untreated countries contribute positive weights.

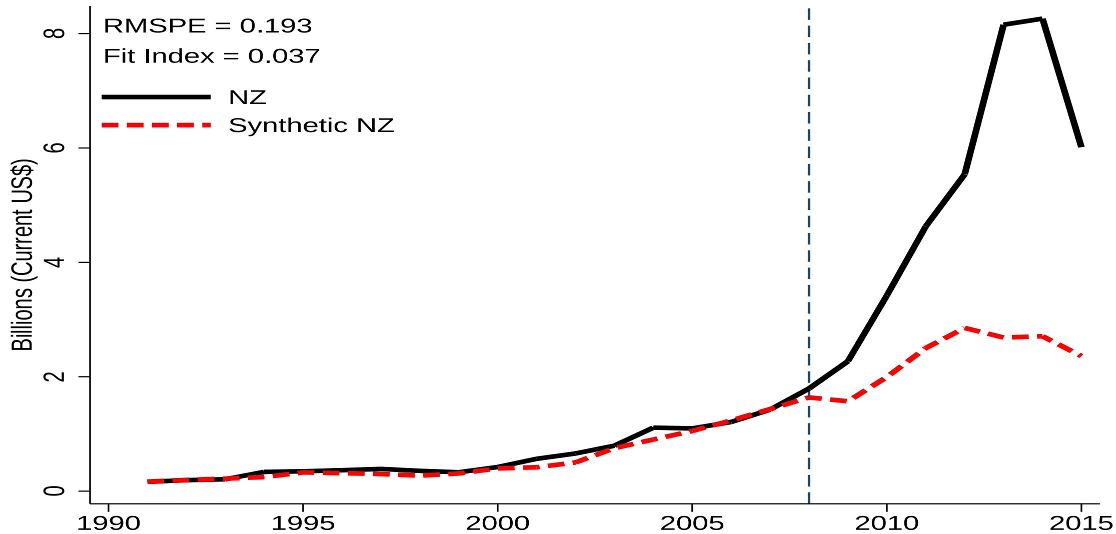

Figure 3: Exports to China: NZ versus synthetic NZ

Next, we perform the placebo tests by iteratively applying the treatment of interest on all other control units within donor pool to test if the estimated impact of the 2008 FTA could be driven entirely by chance. This procedure provides a distribution of estimated placebo treatment effects for all countries within the sample pool. They are placebo effects because the countries in the donor pool, by construction, have never experienced the intervention. We then compare the main synthetic result against all other estimated placebos. If the estimated effect for the NZ case is unusually large compared to the distribution of our control units, we can infer that the treatment of interest had a significant effect. The bold line in panel (a) in Figure 4 shows the synthetic NZ's estimated treatment effect against all other estimated treatment effects from the donor pool. The synthetic NZ stands outside the distribution of the placebo cases, providing some confidence that the estimated effect isn't purely by chance but instead causally related to the FTA.

We make statistical inferences by looking at the ratios between the pre-2008 RMSPE and the post-2008 RMSPE for NZ and for all other countries within the donor pool. The importance of the ratio comes from controlling for a large RMSPE in the post-intervention period, when that same country also has a large RMSPE in the pre-intervention period, thus giving a small post-to-pre ratio. Therefore, a large RMPSE in the post-treatment period does not mean that it is a statistically large effect, if the fit is 'bad' in the pre-treatment period. This post/pre-FTA RMSPE value is shown in panel (b) in Figure 4 and NZ stands out as the country with has highest post/pre-FTA RMSPE. Therefore, if one were to pick a country at random from the sample, the chances of obtaining a ratio as high as observed for NZ is $1/25 = 0.04$, which is the p-value of statistical significance.

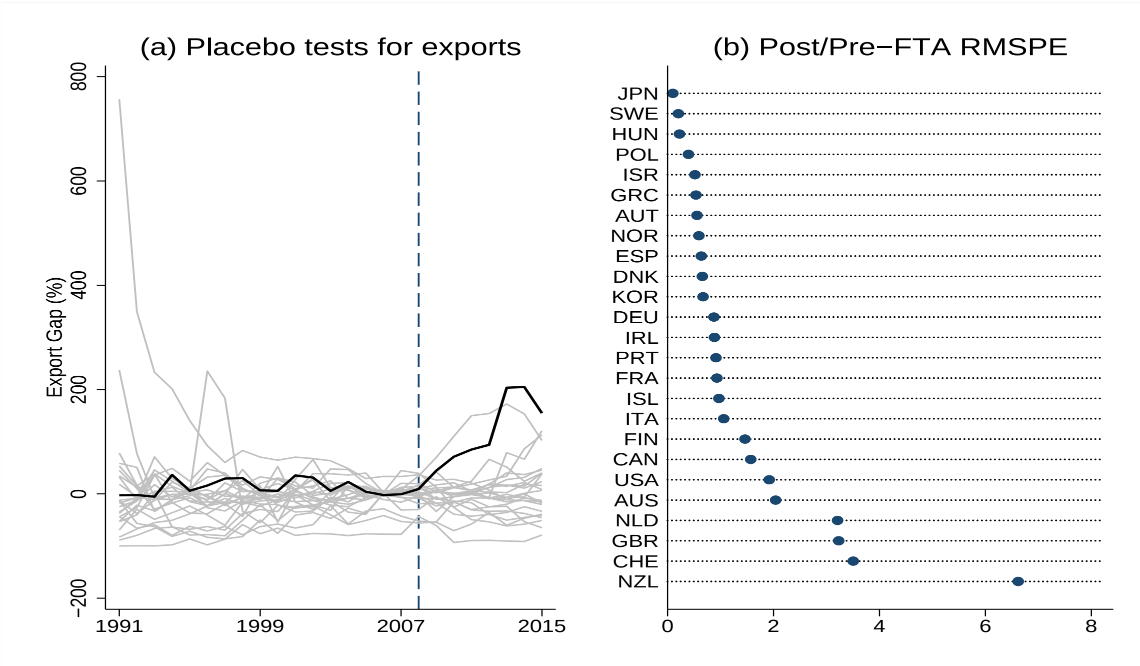

Figure 4: Inference tests for exports

4.3 Exports to the Rest of the World

We employ the SCM to examine the overall effects the 2008 FTA had on NZ's exports to the world and exports to the world excluding China. Keeping the same donor pool and retaining the same covariates that explain trade, we adjust the outcome variable to total exports in current US$. Panel (a) in Figure 5 shows the counterfactual analysis of what NZ's total exports would have been had the FTA never occurred. In the post-intervention period of 2009-2015, actual NZ exports to the rest of the world were 16% higher, on average, than the counterfactual total exports. This finding suggests that huge increases in NZ exports to China in the post-2008 FTA era have outweighed the declining shares of exports going to the other export destinations (such as Australia and the US). In panel (b), the counterfactual analysis of the outcome variable is exports to the rest of the non-Chinese world if NZ never traded with China (i.e., total exports without China). In the post-intervention period of 2009-2015, the actual NZ exports to the rest of the world (excluding China) were 3% higher, on average, than the counterfactual exports. This indicates that if NZ had never signed an FTA with China, total exports to the rest of the non-Chinese world would not have increased more than the actual increases. In order words, total NZ exports to the rest of the world (excluding China) did not decrease in the 2009-2015 period.

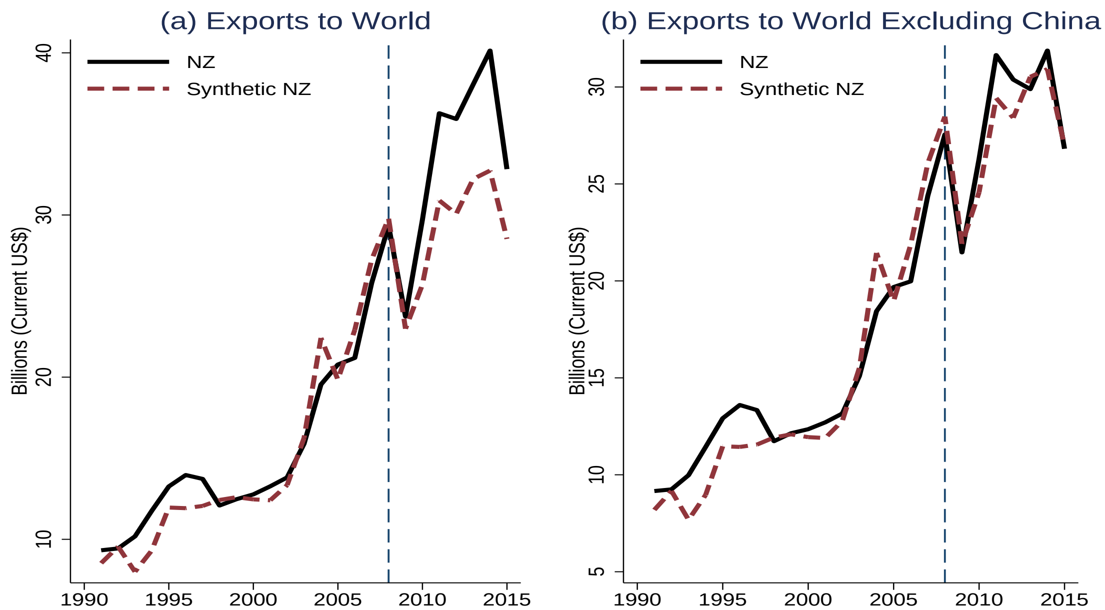

Figure 5: NZ's total exports to the world, including and excluding China

4.4 The Two Most Important Export Sectors

Table 2 shows the NZ industry shares of imports from and exports to China. In 2007, NZ's export share of food and animals (SITC 0) was 36%. By 2017, this share increased to 61.4%, making up more than half of NZ's exports to China. The share of crude oil (SITC 2) being exported to China decreased from 44% in 2007 to 27.4% in 2017. The other remaining products make up only a small share of NZ exports to China. We employ the SCM at the sector level to examine the effects of the NZ-China FTA on NZ's two main exporting industries: (i) food and live animals (SITC 0), and (ii) crude materials (SITC 2).

Panel (a) in Figure 6 shows the evolution of NZ's food and live animals exports, which mainly includes NZ's dairy exports (see Appendix A.2). Had NZ never signed the FTA with China, its food and live animal's exports would have been 130% less, on average, between 2009 and 2015. This suggests that the FTA was an important factor for expanding export growth of NZ's food and live animals sector. The bold line in panel (b) represents the percentage export gap for the food and live animals, the gap between actual exports and its synthetic control, in NZ between 1991 and 2015. The gap for NZ very closely tracks in the zero-gap line, which indicates a good fit. Gray lines represent placebo tests and they are the deviations from synthetic control for the other control countries in the dataset. The next main sector for NZ's exports to China is crude materials (Table 2). Panel (c) suggests that crude materials would have been exporting more than what NZ currently exports if the FTA never been in place. However, the placebo distribution in panel (d) shows that this is not a significant effect due to the large RMSPE in the pre-treatment period, followed by a small post-treatment RMSPE.

Table 2: What does NZ export to and import from China?

		Export share (%)		*Import share (%)*	
SITC	Industry	2007	2017	2007	2017
0	Food and live animals chiefly for food	36.0	61.4	2.3	2.3
1	Beverages and tobacco	0.1	0.5	0.1	0.1
2	Crude materials, inedible, except fuels	44.0	27.4	0.6	0.6
3	Mineral fuels, lubricants and related materials	0.001	0.001	0.1	0.1
4	Animal and vegetable oils, fats and waxes	4.5	0.3	0.1	0.1
5	Chemicals and related products, nes	4.2	2.3	4.9	6.7
6	Manufactured goods classified chiefly by materials	6.0	1.9	16.6	17.1
7	Machinery and transport equipment	4.0	0.8	37.6	42.6
8	Miscellaneous manufactured articles	0.8	0.3	37.7	29.5
9	Commodities and transactions not classified	0.4	5.0	0.2	0.9

Source: United Nations Statistics Division, Commodity Trade Statistics Database.

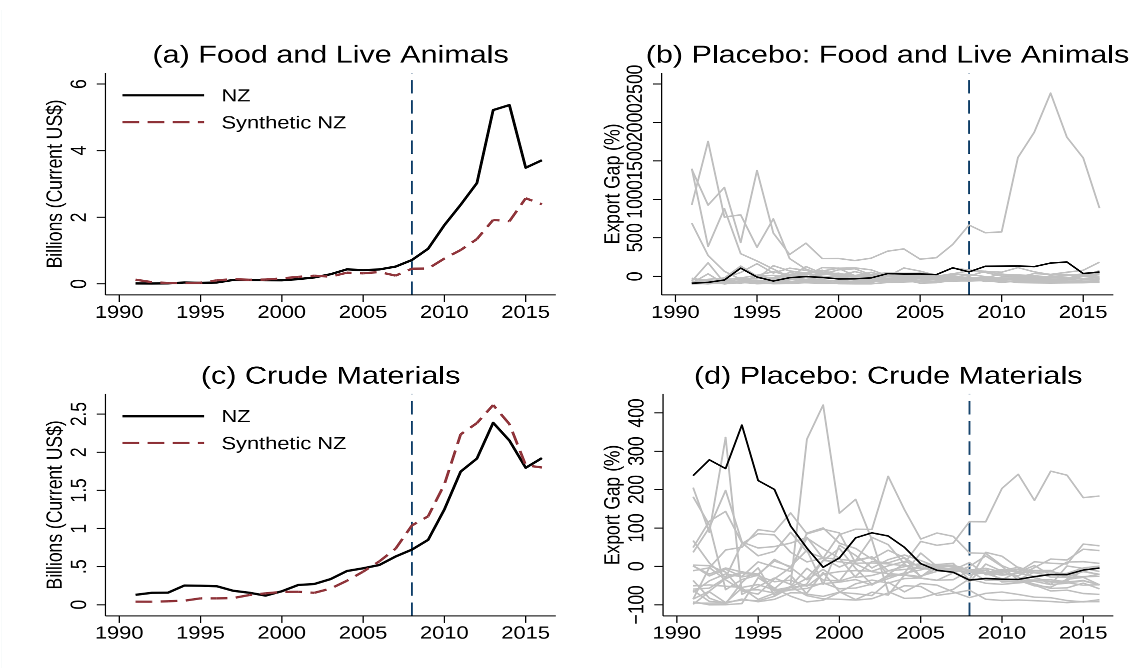

Figure 6: Sector level counterfactuals

5 A Counterfactual Design for Real GDP Per Capita

This section examines the path of real GDP per capita had NZ never signed the FTA with China. We use the latest version of the Penn World Tables' annual country-level data. Our sample period is from 1990 to 2017. The countries we use to construct our synthetic NZ is the same donor pool of 24 countries we introduced in Section 4.1. We use a standard set of economic growth predictors for real per capita GDP, our main dependent outcome variable in this section. This set of covariates are based on a set of growth models within the literature and in a broad sense are meant to capture the impact of institutions, demography and macroeconomic conditions as well as standard growth accounting variables (such as stock of physical and human capital) (Barro, 1991; Mankiw et al., 1992). We also base our set of predictor variables on other SCM models used in the literature, which explore growth effects of structural reforms measured by real GDP per capita (Abadie et al., 2015; Adhikari and Alm, 2016; Meyersson, 2017; Adhikari et al., 2018). The variables we use and data sources are described below.

5.1 Data

Variables from the PWT 9.1: The Penn World Tables (PWT hereafter) provide data on many indicators for a substantial amount of countries on aspects such as relative levels of income, output, production factors, productivity, etc. (Feenstra et al., 2015). The latest version, PWT 9.1, includes 182 countries between 1950 and 2017[19] and features several upgrades in concepts, methods and data sources. From the PWT 9.1, we retain the variables for GDP, population, employment, physical capital, human capital, and openness. Specifically, we use the variable *rgdpe* (expenditure-side real GDP at current PPPs (in millions of 2011 US$)) for GDP, and *pop* (population (in millions)) for population, with which we calculate GDP per capita for each country. The human capital index (variable *hc*) is constructed following the procedure implemented by Hall and Jones (1999) and Caselli (2005), combining years of schooling with the appropriate rates of return. We use the variable *cn* (capital stock at current PPPs (in millions of 2011 US$)) for physical capital stock,[20] and *pop* for population, with which we calculate physical capital per capita for each country. We use the variable *csh_x* (share of merchandise exports at current PPPs) for export share, and *csh_m* (share of merchandise imports at current PPPs) for import share. These are summed (i.e., export share plus import share) to calculate a measure of openness. Finally, we make use of the variable share of population employed. This variable is generated by taking *emp*, which the number of people engaged in working (millions) divided by the total population.

Variables from the WDI: From the WDI (World Bank, 2020), we retain the variables

[19] All versions of the PWT are available at https://www.rug.nl/ggdc/productivity/pwt/.

[20] The physical capital variable (capital stock *cn*) is measured in terms of current PPPs (in millions of 2011 US$) and includes a wide range of assets such as residential and non-residential structures, transport equipment, computers, communication equipment, software, etc.

for age dependency ratio, share of female labor force, fertility, land area, and inflation.[21]

Variables from the Varieties of Democracy (V-Dem): Varieties of Democracy (V-Dem) provides a multidimensional and disaggregated dataset that reflects the complexity of the concept of democracy as a system of rule that goes beyond the simple presence of elections (`https://www.v-dem.net/en/`). From the V-Dem Dataset (Version 9), we retain the variable for democratic participation. There are many measures and varieties for capturing democratic participation of countries. We use the participatory democracy index `v2x_partipdem` put forward by Coppedge et al. (2019). They focus on five key principles that offer a distinctive approach to defining democracy electoral, liberal, participatory, deliberative and egalitarian. In this case the participatory principle of democracy emphasizes active participation by citizens in all political processes. As our donor pool is made up of OECD countries, all countries score high on this index (and on alternative indices).

5.2 Results

Table 3 shows the predictor means for NZ and its synthetic counterpart alongside the OECD average over the pre-intervention period. This demonstrates how well the SCM is performing relative to the OECD average. We calculate the weights of each individual donor country used in the construction of the synthetic counterfactual. From the available number of donor countries, the synthetic control is a weighted combination of France (1.5%), Hungary (43.9%), Israel (30.8%) and Switzerland (23.8%). All other countries obtain zero weights within the donor pool.

Figure 7 plots the time paths of real GDP per capita for NZ and its synthetic counterpart between 1990 and 2017. During the pre-intervention period, the synthetic NZ closely follows actual real GDP per capita. The RMSPE is 0.0282 and the fit index is 0.0063, which suggest that, on average, there is only a 0.16% difference between NZ and its synthetic counterpart in the pre-intervention period.

The estimated impact of the NZ-China FTA on per capita income is given by the difference between the actual and the synthetic real GDP per capita. In the post-intervention era, the 2009-2017 period, the synthetic real GDP per capita was 4%, on average, higher than the actual GDP per capita. In 2009, the synthetic GDP per capita was 4.1% higher

[21] Age dependency ratio is the ratio of dependents – people younger than 15 or older than 64– to the working-age population – those ages 15-64. Female labor force as a percentage of the total shows the extent to which women are active in the labor force. Total fertility rate represents the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with age-specific fertility rates of the specified year. The land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. Inflation is measured by the consumer price index.

than the actual GDP per capita. This continued in the 2009-2015 period, the synthetic real GDP per capita was 4.5%, on average, higher than the actual GDP per capita. The growth of the synthetic GDP per capita, however, slowed down after 2015. In the 2015-2017 period, actual GDP per capita grew at a rate of 3.2% on average, whereas the corresponding figure for the synthetic NZ is 2.0%. Accordingly, in 2017, the synthetic GDP per capita was only 1.6% higher than the actual GDP per capita.

Table 3: GDP per capita predictor means pre-intervention period

	NZ	Synthetic NZ	OECD average
Human capital	3.27	3.19	3.10
Openness	0.56	0.80	0.72
Share of population employed	0.47	0.45	0.46
Capital per capita	65,446	84,695	104,416
Age dependency ratio	52.13	52.29	49.69
Share of female labor force	44.88	44.33	43.58
Fertility	2.01	1.91	1.67
Land area	263,310	63,683	1,267,279
Inflation	2.33	9.24	5.51
Democratic participation	0.70	0.63	0.63
Lagged GDP per capita (1994)	22,972	22,777	25,155
Lagged GDP per capita (1998)	25,959	25,929	30,176
Lagged GDP per capita (2002)	28,370	28,308	32,983
Lagged GDP per capita (2006)	29,576	29,926	37,164
RMSPE		0.0282	

After signing the FTA with China, indicated by the vertical blue dashed line in 2008, NZ's real GDP per capita suffered a small dip, brought on by the 2008-2009 global financial crisis (GFC), but this effect is also captured in the synthetic counterpart: actual GDP per capita contracted by 0.28% in the 2008-2009 period, and the corresponding figure for the synthetic NZ is 0.27%.[22] Figure 7 shows that actual real GDP per capita of NZ and its synthetic counterpart trend upward together in the post-GFC era. In the post GFC recovery period, the 2010-2017 period, actual GDP per capita grew at a rate of 3.17% on average, and the corresponding figure for the synthetic NZ is 2.71%.

[22] The treated country of interest (NZ) should not have undergone any structural shocks during the intervention period, as this would cause doubt on how isolated the treatment effect was. Bilgel and Karahasan (2017) point out structural breaks such as the GFC, that had affected the entire donor pool, does not invalidate the synthetic control estimates.

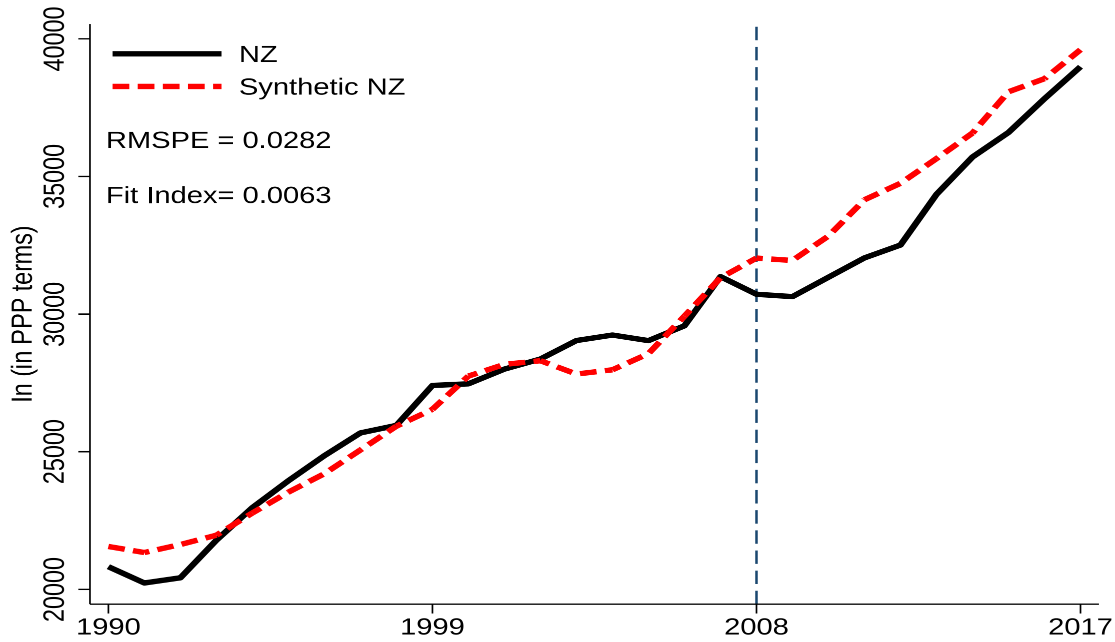

Figure 7: Real GDP per capita counterfactual

5.3 Inference and Pre-Treatment Outcome Lags

Panel (a) in Figure 8 shows the placebo distribution of all other donor countries against our synthetic NZ. After artificially assigning the FTA treatment period on each donor country, we compare where our synthetic NZ sits relative to the distribution and calculate the post/pre value. This post/pre value allows us to identify the exact significance of the estimated effect from the FTA (see panel (b) in Figure 8 for the post/pre values). This ratio is the p-value and can be interpreted as the probability of obtaining a post/pre value that is at least as large as the ones obtained through artificially assigning treatment to the unexposed countries. In other words, if a country had been selected at random from the sample, the probability of obtaining a post/pre RMSPE at least as high as that of NZ is 0.64 (=16/25). This test suggests that the NZ-China FTA had no direct causal effect on NZ's real GDP per capita.

Finally, we conduct a sensitivity test to assess the potential pitfalls that can occur when using the SCM. It has become a popular choice within the application of the SCM to include the entire pre-treatment outcome path of the outcome variable as part of the set of economic predictors (i.e., include all past lagged observations of the dependent variable). Kaul et al. (2017) argue that researchers should never use all pre-intervention outcome periods of the dependent variable as economic predictors. Firstly, they use the Billmeier and Nannicini (2013) data to illustrate the shortcomings of the SCM. More specifically they changed the pre-treatment outcome lags of the dependent variable, in addition to the set of other covariates.

This drastically changed the outcome of the synthetic control and the estimated impact of the treatment effect. Secondly, a further shortcoming of this method is the choice to include the entire pre-treatment outcome path as economic predictors leads to the irrelevancy of other outcome predictors (covariates), which is proven mathematically. We examine the concern of Kaul et al. (2017) that including all outcome lags for pre-treatment years will render the other predictor variables useless and change the outcome trajectory of the synthetic unit. By plotting our original choice of pre-treatment outcome lags against several other combinations of outcome lags, we examine the synthetic unit's sensitivity under different specifications.

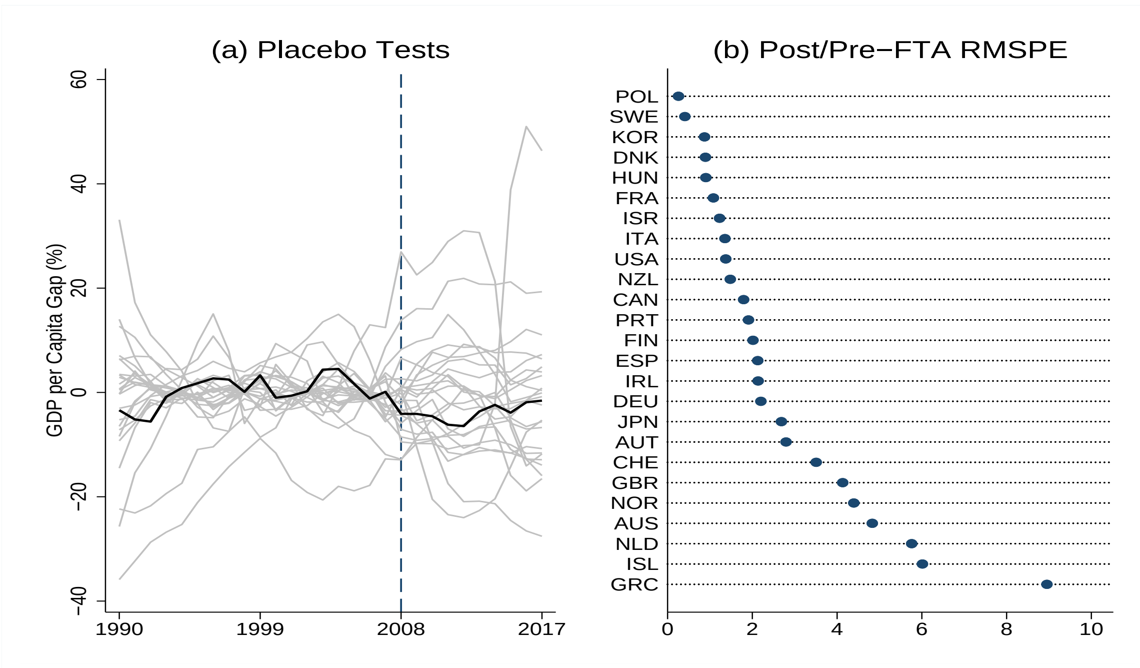

Figure 8: Inference tests

Figure 9 shows the variations in the synthetic outcome due to changes in alternative lagged outcomes. The solid black line shows our actual GDP per capita (in PPP terms) and the solid red line shows the benchmark result that we have provided for our counterfactual analysis. Experiment 1, denoted by 'E1' is GDP per capita with all the past outcome lags. 'E2' includes only even years of the lagged outcome covariate. 'E3' includes outcome lags at 5-year intervals (i.e., 1995, 2000, and 2005). 'E4', following Meyersson (2017), considers only the last 5 years of the pre-intervention period. Following the recommendation of Kaul et al. (2017), 'E5' is just the last outcome lag of the intervention period (in our case it is 2007). All synthetic alternatives follow a similar outcome path to the benchmark specification in the post-intervention period. Average annual growth of the benchmark synthetic real GDP per capita is 2.72% during 2009-2017. All experiments yield similar positive growth rates in the post-intervention period, which range from 2.59% ('E3') to 2.82% ('E4'). In terms of levels of real GDP per capita 'E3' yields the highest levels and 'E5' yields the lowest levels. In 2017, for example, the estimated GDP per capita in 'E3' was 3.4% higher than the benchmark synthetic NZ, and the estimated GDP per capita in 'E5' was 5.9% lower than the benchmark synthetic NZ. The estimated GDP per capita levels in 'E5' were also very low in comparison with the levels of actual and benchmark synthetic GDP per capita in the pre-intervention period, most notably before 2004.

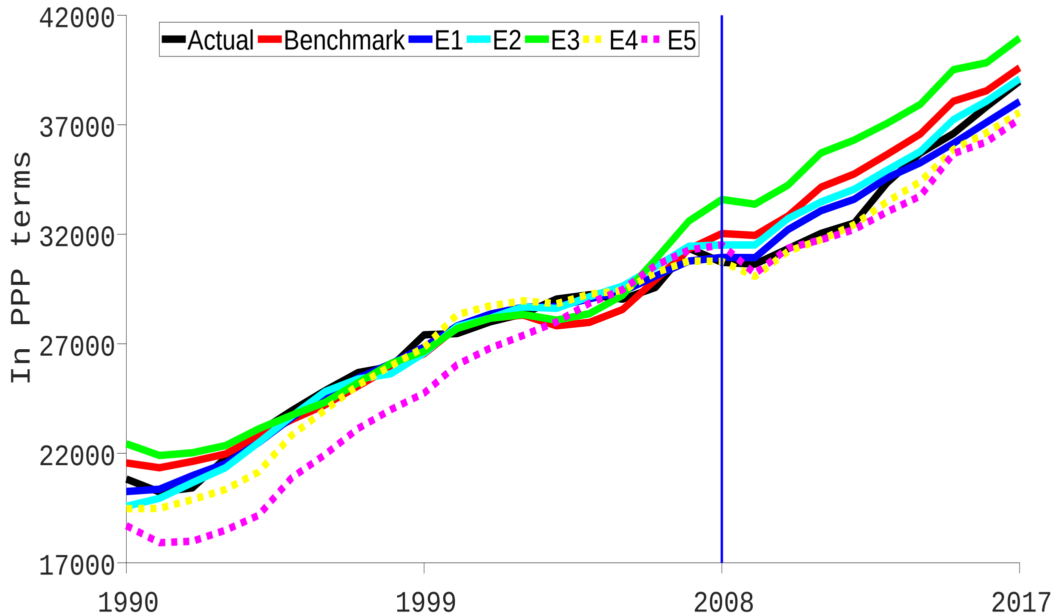

Figure 9: Synthetic NZ GDP per capita with various outcome lags

5.4 Terms of Trade

The choice to use real GDP on the expenditure-side as our measure for economic well-being is in influenced by its inclusion of a country's terms of trade.[23] Terms of trade, the ratio of export prices to import prices, measure a country's purchasing power abroad and is closely related to gains from trade (see Kohli, 2003, 2004; Kehoe and Ruhl, 2008; Reinsdorf, 2010). NZ is a small open economy and relies on its external sectors as a source of economic growth and development. Accordingly, we use the newly developed GDP on the expenditure-side from PWT 9.1 as this accounts for changes in terms of trade.

PWT 9.1 provides two different real GDP measures (at chained PPPs): (i) expenditure-side real GDP to compare relative living standards across countries and over time (variable *rgdpe*); (ii) output-side real GDP to compare relative productive capacity across countries

[23]See Feenstra et al. (2009, p. 201) for a simple example to illustrate this.

and over time (variable *rgdpo*). To demonstrate the difference between these two GDP measures, we plot the two variables for real GDP and per capita real GDP. Panel (a) in Figure 10 shows that GDP series close match each other over time. However, during 1950-2017, *rgdpe* was higher than *rgdpo*. For example, real GDP on the expenditure-side was 3.1% higher than real GDP on the output-side, on average, between 1990 and 2017. A country will have favourable terms of trade if it receives a relatively high price for its exports and pays a relatively low price for its imports. This tends to make *rgdpe* higher than *rgdpo* (Feenstra et al., 2015).[24] This is exactly the case for NZ. We are interested in examining if the FTA with China improved economic welfare for NZ. Accordingly, our results are based on the real GDP on the expenditure-side (variable *rgdpe*).[25]

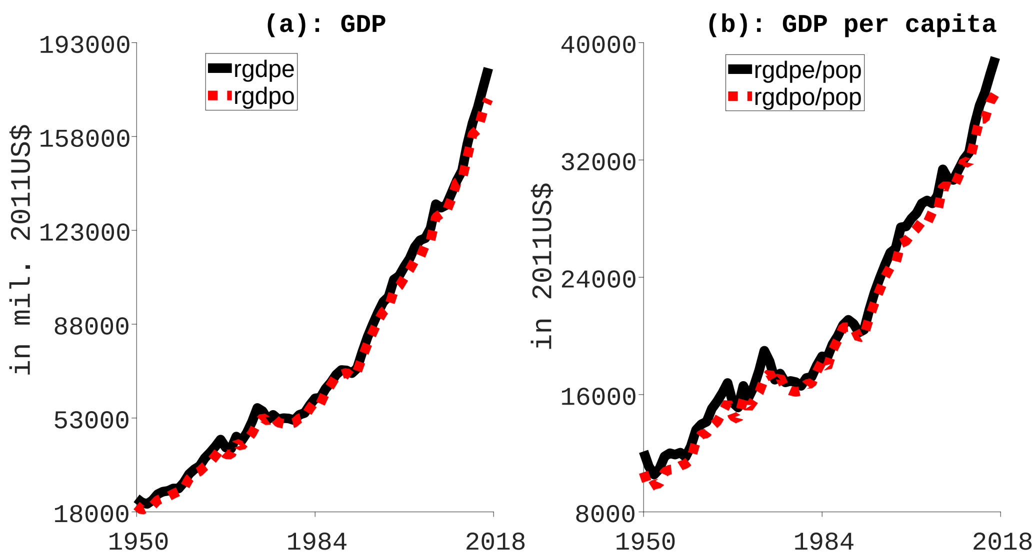

Figure 10: Real GDP^e vs. Real GDP^o at chained PPPs, NZ: 1950-2017

Figure 11 shows the evolution of terms of trade in NZ between 2000 and 2018.[26] NZ's terms of trade increased markedly between 2000 and 2007. NZ is a primary commodity exporter, and the increase in the country's terms of trade during the 2000s reflected rising export prices. Declines in the terms of trade occurred over the GFC and the huge reduction in dairy prices of 2014 is reflected in Figure A4. This discussion adds to the idea that by utilizing the real GDP per capita from the PWT 9.1, we are able to capture an important driving factor for NZ's economic growth performance.

[24] The gap between these two measures of real GDP is not a measure of the gains from trade for countries, or at least, not the gains from trade as compared to autarky.

[25] We replicate our exercises using the variable *rgdpo* and find the results reassuringly robust.

[26] We use the 'net barter terms of trade index (2000=100)' series from the World Bank (2020).

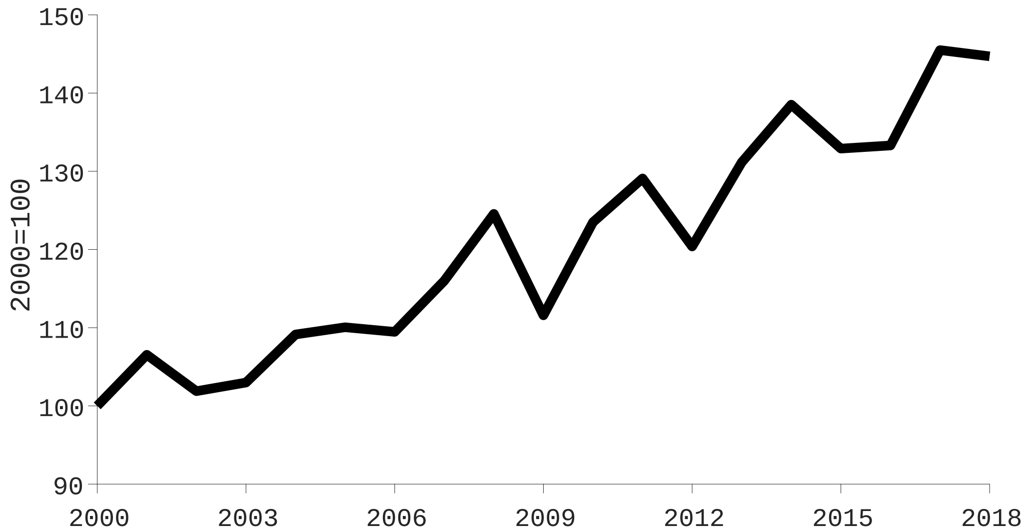

Figure 11: NZ's terms of trade, 2000-2018

6 Some Robustness Exercises

Our two main results in this paper are presented in Figure 3 (for NZ's exports to China) and Figure 7 (for NZ's real GDP per capita). Figure 3 presents clear gains of the 2008 NZ-China FTA for NZ's exports to China. There is a very sizeable difference between the actual and the synthetic NZ's exports to China in the post-treatment period. However, we do not observe such a substantial difference between the actual and the synthetic real GDP per capita series in Figure 7. In fact, actual GDP per capita lies below the synthetic GDP per capita during the 2009-2017 period, suggesting negative impacts on NZ's income per capita due to the FTA. In this section we provide two robustness exercises to examine to what extent this negative effect is robust.

6.1 Extending the Pre-Treatment Sample Period

Our benchmark specification for real GDP per capita results is based on the 1990-2017 sample period. In our first exercise, we extend the pre-treatment period to explore the possible effects of a longer pre-treatment period. We study the 1977-2017 period with the same 24 donor countries. We employ the same variables, except the share of female labor force due to the lack of data. Figure 12 plots the time paths of real GDP per capita for NZ and its synthetic counterpart between 1977 and 2017. Figure 12 presents a very similar picture to our benchmark results presented in Figure 7, both in terms of pre- and post-treatment paths. In comparison with our benchmark results presented in Figure 7, Figure 12 reports a higher RMSPE (0.0444 versus 0.0282) and fit index (0.0073 versus 0.0063), although the reported statistics are not too far from our benchmark statistics. Nevertheless, our benchmark specification has a better fit. In the post-treatment period, a comparison of Figure 7 and Figure 12 suggests that the generated synthetic GDP per capita series very similar. According to our benchmark results, the synthetic GDP per capita was 4%, on average, higher than the actual GDP per capita during 2009-2017. The corresponding value based on the experiment presented in Figure 12 is 4.2%

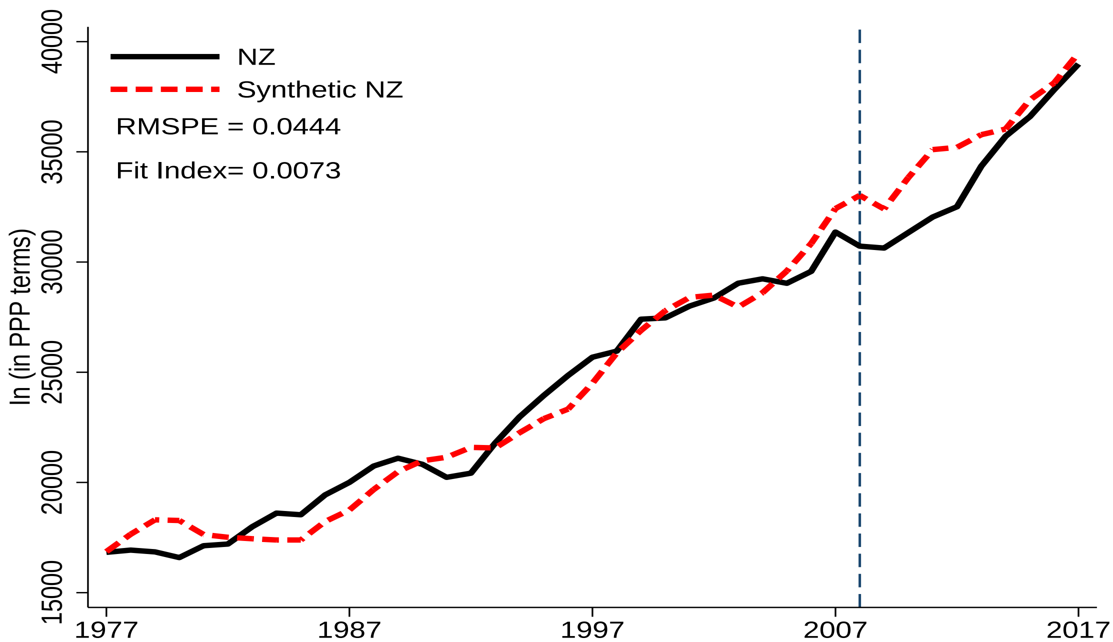

Figure 12: Real GDP per capita counterfactual, 1977-2017

6.2 Changing the Control Group

In our second exercise, we consider that some of the European countries in our control group went through the process of currency union. Although all EU countries are part of the Economic and Monetary Union (EMU), 19 of them have replaced their national currencies with the single currency the euro. These EU countries form the euro area, also known as the Eurozone.[27] We drop Austria, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, and Spain, a total of 10 countries from our original 24 donor countries. We retain all the same variables as in our benchmark case. In other words, we only change the control group. Figure 13 plots the time paths of real GDP per capita for NZ and its synthetic counterpart between 1990 and 2017, excluding the 10 Eurozone countries from our benchmark control group. In comparison with our benchmark results presented

[27] https://europa.eu/european-union/about-eu/euro/which-countries-use-euro_en

in Figure 7, Figure 13 reports a lower RMSPE (0.0198 versus 0.0282) and fit index (0.0046 versus 0.0063), indicating a better fit than the benchmark specification. The euro is the second most-used currency worldwide and is a solid indicator of European integration. NZ is not a member of this integration and this may explain the better pre-treatment period of the country group without the 10 members of the EMU.

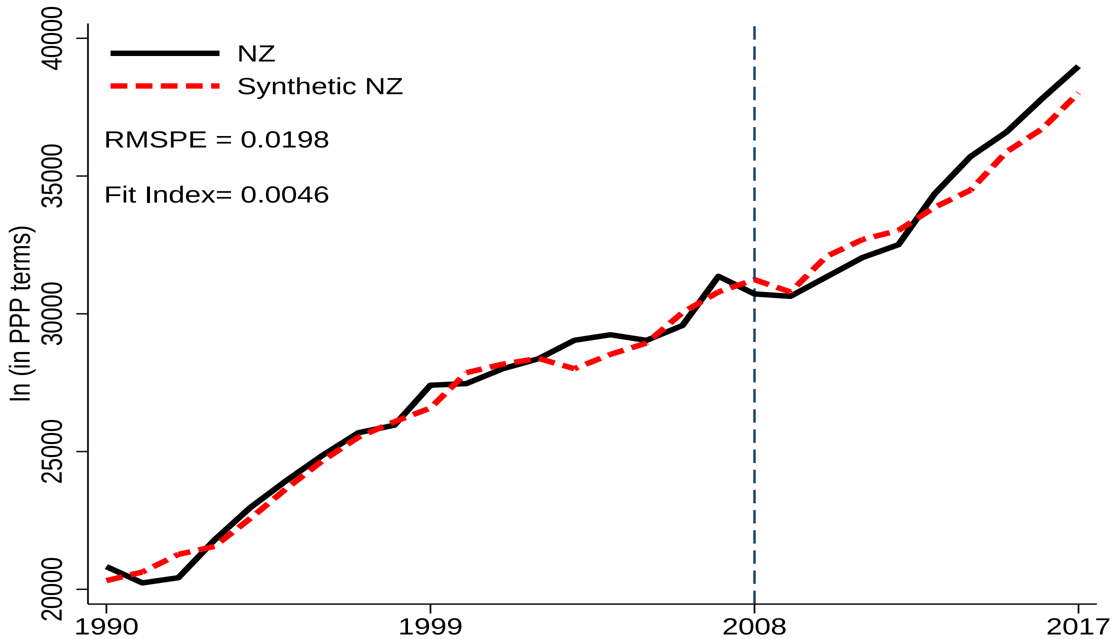

Figure 13: Real GDP per capita counterfactual, without the Eurozone countries

The synthetic per capita GDP presented in Figure 13 lies above the actual one between 2009 and 2012, and below the actual GDP per capita between 2013 and 2017. The behavior of the synthetic GDP per capita in the post-intervention period is different to our benchmark result presented in Figure 7. According to our benchmark results, the 2008 NZ-China FTA did not cause an increase in NZ's real GDP per capita, because the synthetic GDP per capita lies above the actual GDP per capita in the post-intervention period of the 2009-2017. Our experiment without the Eurozone countries yields the same result between 2009 and 2012, where the actual GDP per capita lies below than the synthetic GDP per capita in Figure 13. However, the experiment without the Eurozone countries suggests gains in real GDP per capita starting with the year 2013, where the actual GDP per capita was 2.5%, on average, higher than the synthetic GDP per capita during 2013-2017.

We also re-examine the time path of the synthetic NZ's exports to China using the country group without the Eurozone countries. Figure 14 presents the actual NZ exports to China, our benchmark counterfactual (based on the 24-country control group), and the counterfactual we estimate for robustness (the control group excluding Eurozone countries).

Both counterfactuals behave similarly in the pre-treatment period, i.e., it is very difficult to distinguish the solid blue line from the solid red line over the 1991-2007 period. In the post-FTA era, both series follow a similar outcome path well below NZ's actual exports to China. Both counterfactuals suggest that NZ's exports to China increased significantly due to the 2008 FTA. Although both counterfactuals demonstrate the same qualitative effect, there are differences in the size of the effects. The solid red line lies below the solid blue line in the post-treatment period, suggesting our benchmark counterfactual implies larger gains from the FTA. For example, in 2014, NZ's export to China amounted US$ 8.26 billion. Our benchmark counterfactual suggests that the corresponding figure would have been US$ 2.71 billion in 2014, which yields an increase of US$ 5.55 billion in NZ's export China due to the FTA. The experiment without the Eurozone countries suggests that NZ's exports to China would have been US$ 3.62 billion in 2014, which yields an increase of US$ 4.64 billion in NZ's export China due to the FTA.

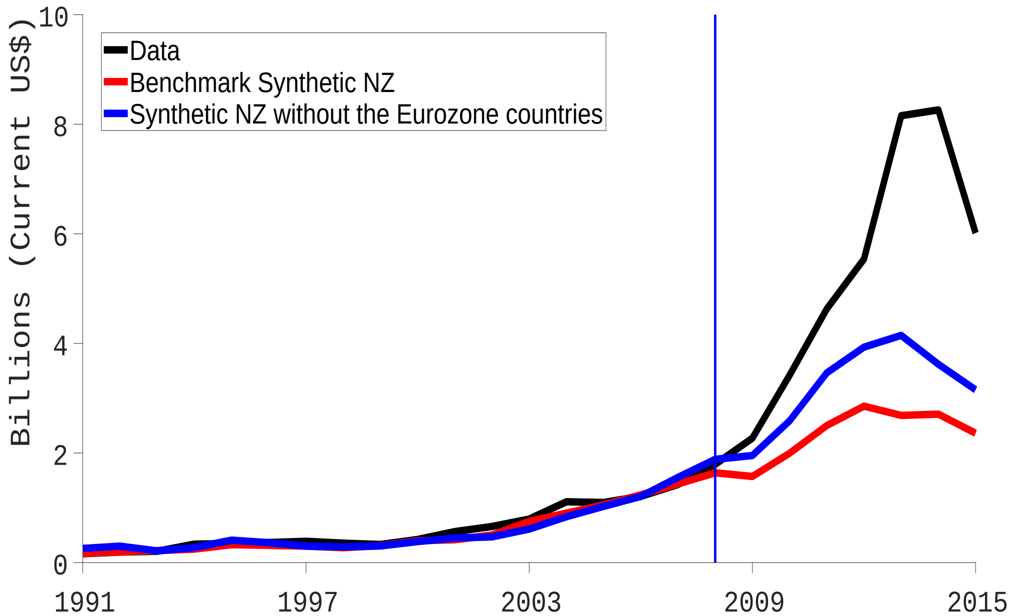

Figure 14: Exports to China: Effects of the control group

7 Conclusions

Our paper is timely as FTAs all over the world are being negotiated or renegotiated. We highlight that there is evidence to suggest that the 2008 NZ-China FTA has been an important factor for NZ's export growth. While we find substantial export gains, owing to sectors such as dairy products, we do not find significant impact on GDP per capita of NZ. Both the benchmark and the robustness cases indicate a small but negative effect of the FTA on NZ's real GDP per capita between 2009 and 2012. This negative effect loses its magnitude in the 2013-2017 period in the benchmark case (where we have 24 countries in our control group); and it becomes positive in the robustness case (where we exclude the Eurozone countries from the control group). It is important to note that real income per capita tends to be a much slower moving variable than exports. Hence, a much longer post-intervention period can clarify the benefits of the 2008 FTA on NZ's per capita income. Our post-intervention period covers the 2009-2017 period for NZ's real GDP per capita, which is less than a decade. Replicating our study for the 2009-2027 or 2009-2037 periods, for example, would be much more informative regarding the long-term benefits of the 2008 NZ-China FTA for the NZ economy. This is our suggestion for future research for the upcoming decades. In addition, domestic policy settings are likely impacting the distributional effects that expanding export growth are expected to have on income, which could act as an offsetting force.

Although not all industries are examined here, our analysis shows that the dairy sector, which is NZ's primary exporting sector, incurred a substantial gain from the NZ-China FTA. The caveat to these findings is the concentration shift at the sector level of NZ's exports to China. This is especially important within the dairy sector, whose exports constitute more than half of total NZ exports to China.[28] China is NZ's biggest export partner in terms of merchandise goods, and any predicted downturn in the Chinese economy would have implications for NZ's trade flows. In this respect, the future developments in the Comprehensive and Progressive Trans-Pacific Partnership may provide NZ with an opportunity to further diversify its export partners and insulated itself from any unexpected exogenous shocks in prices and overseas demand.[29]

References

Abadie, A., Diamond, A., Hainmueller, J. 2015. Comparative politics and the synthetic control method. American Journal of Political Science, 59(2), 495-510.

[28] China's recent urbanization is a driver for the huge increase in demand for NZ milk powder, which dominates the NZ dairy exports to China (Kendall, 2014).

[29] The Trans-Pacific Partnership (TPP) was a FTA designed to liberalize trade and investment between 12 Pacific-rim countries: NZ, Australia, Brunei Darussalam, Canada, Chile, Japan, Malaysia, Mexico, Peru, Singapore, the US and Vietnam. The concluded agreement was signed on 4 February 2016, and NZ ratified this agreement in May 2017. When the US dropped out of the TPP, remaining 11 countries signed an amended agreement on 8 March 2018 in Chile. The TPP is now known as the Comprehensive and Progressive Trans-Pacific Partnership (CPTPP). NZ ratified the CPTPP on 25 October 2018. With Australia's ratification on 30 December 2018, the CPTPP has met the threshold requirements to enter into force.

Abadie, A., Diamond, A., Hainmueller, J. 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.

Abadie, A., Cattaneo, M. D. 2018. Econometric methods for program evaluation. Annual Review of Economics, 10, 465-503.

Abadie, A., L'Hour, J. 2019. A penalized synthetic control estimator for disaggregated data. Unpublished manuscript, `https://economics.mit.edu/faculty/abadie/papers` (13 November 2019 version).

Abadie, A., Gardeazabal, J. 2003. The economic costs of conflict: A case study of the Basque Country. American Economic Review, 93(1), 113-132.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., Mitton, T. 2016. The value of connections in turbulent times: Evidence from the United States. Journal of Financial Economics, 121(2), 368-391.

Adhikari, B., Alm, J. 2016. Evaluating the economic effects of flat tax reforms using synthetic control methods. Southern Economic Journal, 83(2), 437-463.

Adhikari, B., Duval, R., Hu, B., Loungani, P. 2018. Can reform waves turn the tide? Some case studies using the synthetic control method. Open Economies Review, 29(4), 879-910.

Amjad, M., Shah, D., Shen, D. 2018. Robust synthetic control. Journal of Machine Learning Research, 19, 1-51.

Athey, S., Imbens, G. W. 2017. The state of applied econometrics: Causality and policy evaluation. Journal of Economic Perspectives, 31(2), 3-32.

Aytuğ, H., Kütük, M. M., Oduncu, A., Togan, S. 2017. Twenty years of the EU-Turkey Customs Union: A synthetic control method analysis. Journal of Common Market Studies, 55(3), 419-431.

Barone, G., Mocetti, S. 2014. Natural disasters, growth and institutions: A tale of two earthquakes. Journal of Urban Economics, 84, 52-66.

Barro, R. J. 1991. Economic growth in a cross section of countries. Quarterly Journal of Economics, 106(2), 407-443.

Becker, M., Klößner, S. 2018. Fast and reliable computation of generalized synthetic controls. Econometrics and Statistics, 5, 1-19.

Bertrand, M., Duflo, E., Mullainathan, S. 2004. How much should we trust differences-in-differences estimates? Quarterly Journal of Economics, 119(1), 249-275.

Bilgel, F., Galle, B. 2015. Financial incentives for kidney donation: A comparative case study using synthetic controls. Journal of Health Economics, 43, 103-117.

Bilgel, F., Karahasan, B. C. 2017. The economic costs of separatist terrorism in Turkey. Journal of Conflict Resolution, 61(2), 457-479.

Billmeier, A., Nannicini, T. 2013. Assessing economic liberalization episodes: A synthetic control approach. Review of Economics and Statistics, 95(3), 983-1001.

Bove, V., Elia, L. 2018. Economic development in peacekeeping host countries. CESifo Economic Studies, 64(4), 712-728.

Bove, V., Elia, L., Smith, R. P. 2017. On the heterogenous consequences of civil war. Oxford Economic Papers, 69(3), 550-568.

Bowman, S., Conway, P. 2013a. China's recent growth and its impact on the New Zealand Economy. New Zealand Treasury Working Paper 13/15.

Bowman, S., Conway, P. 2013b. The outlook for China's growth and its impact on the New Zealand exports. New Zealand Treasury Working Paper 13/16.

Caselli, F. 2005. Accounting for cross-country income differences. In: Aghion, P., Durlauf, S. (Eds.), Handbook of Economic Growth, Elsevier Press, pp. 679-741.

Cavallo, E., Galiani, S., Noy, I., Pantano, I. 2013. Catastrophic natural disasters and economic growth. Review of Economics and Statistics, 95(5), 1549-1561.

Chung, S., Lee, J., Osang, T. 2016. Did China tire safeguard save U.S. workers? European Economic Review, 85, 22-38.

Coffman, M., Noy, I. 2012. Hurricane Iniki: Measuring the long-term economic impact of a natural disaster using synthetic control. Environment and Development Economics, 17(2), 187-205.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K. L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Wilson, S., Cornell, A., Gastaldi, L., Gjerløw, H., Ilchenko, N., Krusell, J., Maxwell, L., Mechkova, V., Medzihorsky, J., Pernes, J., von Römer, J., Stepanova, N., Sundström, A., Tzelgov, E., Wang, Y.-t., Wig, T., Ziblatt, D. 2019. V-Dem [Country-Year/Country-Date] Dataset v9, Varieties of Democracy (V-Dem) Project, `https://doi.org/10.23696/vdemcy19`

Costalli, S., Moretti, L., Pischedda, C. 2017. The economic costs of civil war: Synthetic counterfactual evidence and the effects of ethnic fractionalization. Journal of Peace Research, 54(1), 80-94.

Doudchenko, N., Imbens, G. W. 2016. Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. NBER Working Paper No. 22791.

Feenstra, R. C., Heston, A., Timmer, M. P., Deng, H. 2009. Estimating real production and expenditures across nations: A proposal for improving the Penn World Tables. Review of Economics and Statistics, 91(1), 201-212.

Feenstra, R. C., Inklaar, R., Timmer, M. P. 2015. The next generation of the Penn World Table. American Economic Review, 105(10), 3150-3182.

Ferman, B., Pinto, C. 2017. Placebo tests for synthetic controls. Unpublished manuscript, `https://sites.google.com/site/brunoferman/research` (March 2017 version).

Ferman, B., Pinto, C., Possebom, V. 2018. Cherry picking with synthetic controls. Unpublished manuscript, `https://sites.google.com/site/brunoferman/research` (March 2018 version).

Firpo, S., Possebom, V. 2018. Synthetic control method: Inference, sensitivity analysis and confidence sets. Journal of Causal Inference, 6(2), 1-26.

Gardeazabal, J., Vega-Bayo, A. 2017. An empirical comparison between the synthetic control method and Hsiao et al.'s panel data approach to program evaluation. Journal of Applied Econometrics, 32(5), 983-1002.

Gharehgozli, O. 2017. An estimation of the economic sost of recent sanctions on Iran using the synthetic control method. Economics Letters, 157, 141-144.

Grier, K., Maynard, N. 2016. The economic consequences of Hugo Chavez: A synthetic control analysis. Journal of Economic Behavior & Organization, 125, 1-21.

Hahn, J., Shi, R. 2017. Synthetic control and inference. Econometrics, 5(4), 1-12.

Hall, R. E., Jones, C. I. 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics, 114(1), 83-116.

Hannan, S. A. 2017. The impact of trade agreements in Latin America using the synthetic control method. International Monetary Fund Working Paper No. WP/17/45.

Hannan, S. A. 2016. The impact of trade agreements: New approach, new insights. International Monetary Fund Working Paper No. WP/16/117.

Head, K., Mayer, T. 2014. Gravity equations: Workhorse, toolkit, and cookbook, in: Gopinath, G., Helpman, E., and Rogoff, K. Handbook of International Economics Volume 4, Elsevier Press, pp. 131-195.

Holland, P. W. 1986. Statistics and causal inference. Journal of the American Statistical Association, 81(396), 945-960.

Hosny, A. S. 2012. Algeria's trade with GAFTA countries: A synthetic control approach. Transition Studies Review, 19(1), 35-42.

Imbens, G. W., Rubin, D. B. 2015. Causal inference for statistics, social, and biomedical sciences: An introduction. Cambridge: Cambridge University Press.

Imbens, G. W., Wooldridge, J. M. 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature, 47(1), 5-86.

Jales, H., Kang, T. H., Stein, G., Ribeiro, F. G. 2018. Measuring the role of the 1959 revolution on Cuba's economic performance. The World Economy, 41(8), 2243-2274.

Kang, Y., Beal, T. 2013. Barriers to New Zealand-China economic integration: A case of the dairy industry and beyond. New Zealand Journal of Asian Studies, 15(1), 1-17.

Kaul, A., Klößner, S., Pfeifer, G., Schieler, M. 2017. Synthetic control methods: Never use all pre-intervention outcomes together with covariates. Unpublished Manuscript (29 July 2017 version), https://mpra.ub.uni-muenchen.de/83790/1/MPRA_paper_83790.pdf

Kehoe, T. J., Ruhl, K. J. 2008. Are shocks to the terms of trade shocks to productivity? Review of Economic Dynamics, 11(4), 804-819.

Kendall, R. 2014. Economic linkages between New Zealand and China. Reserve Bank of New Zealand Analytical Notes AN2014/06.

Kohli, U. 2004. Real GDP, real domestic income, and terms-of-trade changes. Journal of International Economics, 62(1), 83-106.

Kohli, U. 2003. Terms of trade, real GDP, and real value added: A new look at New Zealand's growth performance. New Zealand Economic Papers, 37(1), 41-66.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., Sutton, M. 2016. Examination of the synthetic control method for evaluating health policies with multiple treated units. Health Economics, 25(12), 1514-1528.

Mankiw, N. G., Romer, D., Weil, D. N. 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics, 107(2), 407-437.

Matta, S., Appleton, S., Bleaney, M. 2019. The impact of the Arab spring on the Tunisian economy. The World Bank Economic Review, 33(1), 231-258.

Meyersson, E. 2017. Pious populists at the gate: A case study of economic development in Turkey under AKP. Economics of Transition, 25(2), 271-312.

Mideksa, T. K. 2013. The economic impact of natural resources. Journal of Environmental Economics and Management, 65(2), 277-289.

Ministry of Foreign Affairs and Trade (MFAT). 2007. New Zealand-China free trade agreement, national interest analysis. Retrieved from: https://www.mfat.govt.nz/assets/FTAs-agreements-in-force/China-FTA/National-interest-analysis.pdf

Mohan, P. 2017. The economic impact of hurricanes on bananas: A case study of Dominica using synthetic control methods. Food Policy, 68, 21-30.

Montalvo, J. G. 2011. Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. Review of Economics and Statistics, 93(4), 1146-1154.

Nannicini, T., Billmeier, A. 2011. Economies in transition: How important is trade openness for growth? Oxford Bulletin of Economics and Statistics, 73(3), 287-314.

Olper, A., Curzi, D., Swinnen, J. 2018. Trade liberalization and child mortality: A synthetic control method. World Development, 110, 394-410.

Osborn, D. R., Vehbi, T. 2013. Empirical evidence on growth spillovers from China to New Zealand. New Zealand Treasury Working Paper 13/17.

Pinotti, P. 2015. The economic costs of organised crime: Evidence from Southern Italy. Economic Journal, 125(586), F203-F232.

Powell, B., Clark, J. R., Nowrasteh, A. 2017. Does mass immigration destroy institutions? 1990s Israel as a natural experiment. Journal of Economic Behavior & Organization, 141, 83-95.

Puzzello, L., Gomis-Porqueras, P. 2018. Winners and losers from the €uro. European Economic Review, 108, 129-152.

Reinsdorf, M. B. 2010. Terms of trade effects: Theory and measurement. Review of Income and Wealth, 56(s1), S117-S205.

Ritzel, C., Kohler, A. 2017. Protectionism, how stupid is this? The causal effect of free trade for the world's poorest countries: Evidence from a quasi-experiment in Switzerland. Journal of Policy Modeling, 39(6), 1007-1018.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 66(5), 688-701.

Saia, A. 2017. Choosing the open sea: The cost to the UK of staying out of the euro. Journal of International Economics, 108, 82-98.

Singhal, S., Nilakantan, R. 2016. The economic effects of a counterinsurgency policy in India: A synthetic control analysis. European Journal of Political Economy, 45, 1-17.

Smith, B. 2015. The resource curse exorcised: Evidence from a panel of countries. Journal of Development Economics, 116, 57-73.

Stojkov, K., Noy, I., Sağlam, Y. 2018. The trade impacts of a food scarce: The Fonterra contamination incident. Journal of Agricultural & Food Industrial Organization, 16(2), 1-16.

Tinbergen, J. 1962. Shaping the World economy: Suggestions for an international economic policy. New York: Twentieth Century Fund.

Verevis, S. 2018. 2008 New Zealand-China free trade agreement: A quantitative analysis. Unpublished Master Thesis. University of Otago.

World Bank. 2018. World Development Indicators. `http://databank.worldbank.org/ddp/home.do`. Accessed on February-December 2018.

World Bank. 2020. World Development Indicators. http://databank.worldbank.org/ddp/home.do. Accessed on January 2020.

Xu, Y. 2017. Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25(1), 57-76.

Appendix A

Appendix A.1 Gravity Features of Trade Data

The graphical representations of trade data show the key gravity features of bilateral trade flows and motivate our use of the gravity model (see also Head and Mayer, 2014, Section 1.1). The first important feature of trade data is how exports and imports rise proportionately with the economic size of a country. Using GDP as a proxy for the economy size, Figure A1 shows NZ's exports (imports) to (from) OECD countries (and China) in 2006.[30] On the vertical axis is NZ's exports to (or imports from) OECD countries (and China) and on the horizontal axis is GDP (in current US dollars) of the trade partner for 2006. The line shows the predicted values for a simple regression of log trade flows on log GDP. Both panels in Figure A1 suggest a near unit elasticity which demonstrate that trade flows rise in proportion to the size of the economic destination. Figure A2 illustrates the second key empirical relationship embodied in the gravity equation, the strong negative relationship that exists between physical distance and trade.

Appendix A.2 A Look at the Dairy Sector

Figure A3 clearly shows that dairy products and birds' eggs (SITC 02) is the largest component of food and live animals (SITC 0) exports to China (SITC 02/SITC 0). Figure A3 also plots the share of SITC 02 in total exports going to China (SITC 02/SITC Total) between 2000 and 2018, and suggest that dairy and birds' eggs is a major component of exports to China (SITC 02/SITC Total). There is a noticeably sharp increase in NZ's dairy exports to China in the post-2008 NZ-China FTA era. These sharp increases in dairy exports may be driven by reduced tariffs and change in food/dairy prices. Figure A4 shows the evolution of the FAO food price index (FFPI) and the FAO dairy price index (FDPI) between 1990 and 2018 (where the base period is 2002-2004) to explore this point further.[31] Food prices (and dairy prices) experienced a strong upswing in 2007 and 2008, reaching historically high levels. For the whole of 2008, the FFPI averaged 201.4 points, up 24.8% from 2007 and representing the highest annual average. Food prices declined 18.7% from 2014 to 2015. The decline was more severe for dairy price index. For the whole of 2015, the FDPI averaged

[30] For Figure A1-A2 we combine the countries Belgium and Luxembourg, which is labelled BLX.

[31] Data are available at http://www.fao.org/worldfoodsituation/foodpricesindex/en/. The Food and Agriculture Organization (FAO) FFPI was introduced in 1996. The only major modification made to it was in 2009, when its base period was updated to 2002-2004. FFPI consists of the average of five commodity group price indices (meat, dairy, cereals, vegetable oil, and sugar), weighted with the average export shares of each of the groups for 2002-2004. FDPI consists of butter, skim milk powder, whole milk powder, and cheese price quotations and world average export trade shares weight the average for 2002-2004.

160.3 points, down 28.5% from 2014. This also provides an explanation behind the large drop in NZ exports to China during 2014-2015, which coincides with a drop in the FFPI and the FDPI. Table 2 displays that more than half of the NZ exports to China is the category of food and live animals (SITC 0). Therefore, a drop in the dairy price index is significant contributing factor to the decline in NZ total exports to China over 2014-2015.

Lastly, we mention the so-called Fonterra contamination incident. On 3 August 2013, Fonterra, one of the world's largest dairy companies, informed the NZ's Ministry for Primary Industries (MPI) of a possible food safety issue relating to a suspected contamination of whey protein concentrate with Clostridium botulinum. China ordered the recall of potentially contaminated milk imported from NZ after this incident. This temporary ban didn't impact whole milk powder and skim milk powder or imports of infant milk formula that had already been processed.[32] Stojkov et al. (2018) note that the affected products made up only a small proportion of NZ's dairy exports, with the vast majority of dairy exports being unaffected. On 28 August 2013, MPI announced that further testing confirmed no products had been contaminated. The incident was a false alarm. However, the Chinese ban remained in place until October 2014. Stojkov et al. (2018) examine the impact of a food contamination scare on dairy exports using the SCM. They find that there was an initial negative shock to the exports of products that were thought to have been contaminated, but that there were no significant sustained impacts on other dairy products. They argue that total dairy exports were 0.6-1.9% lower than they would have been had the scare not occurred.

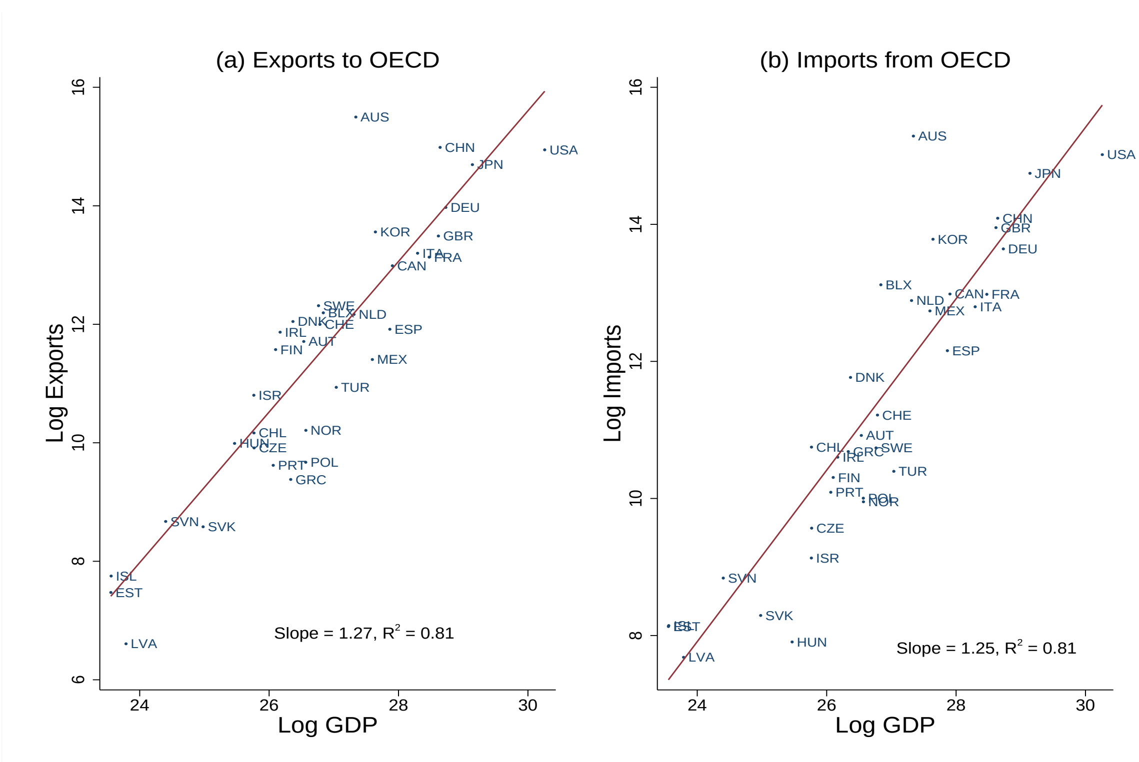

Figure A1: NZ's exports to and imports from OECD (and China), 2006

[32]https://www.nbr.co.nz/article/china-lifts-temporary-ban-fonterra-whey-products-one-russia-remains-bd-164795

Figure A2: Trade inversely proportional to distance, 2006

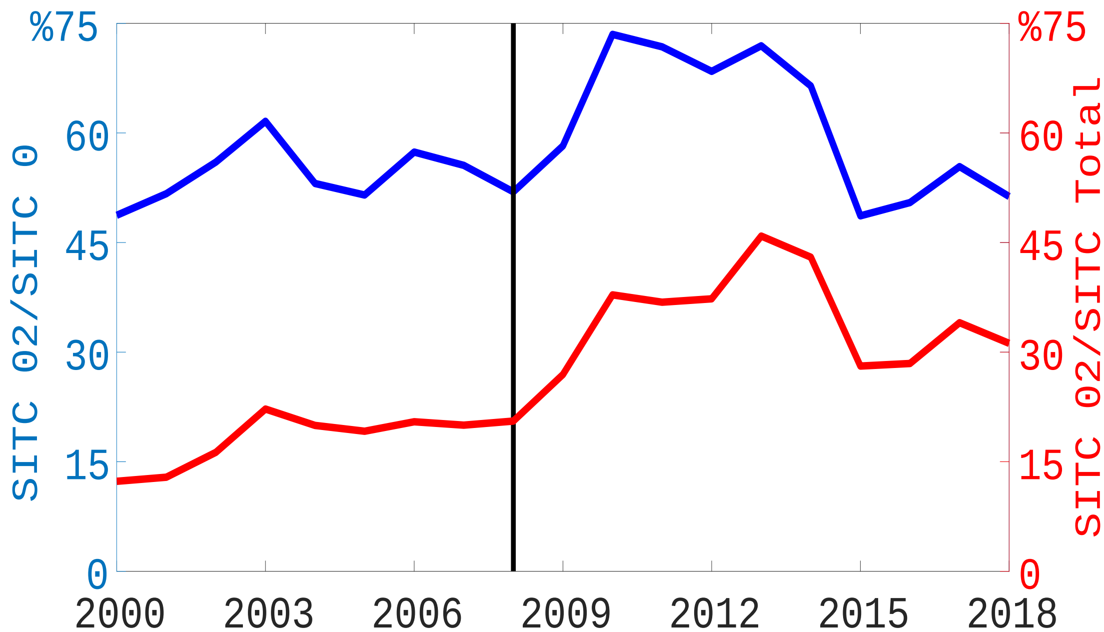

Figure A3: NZ's *dairy products* exports to China (%), 2000-2018

Source: `https://comtrade.un.org/data`

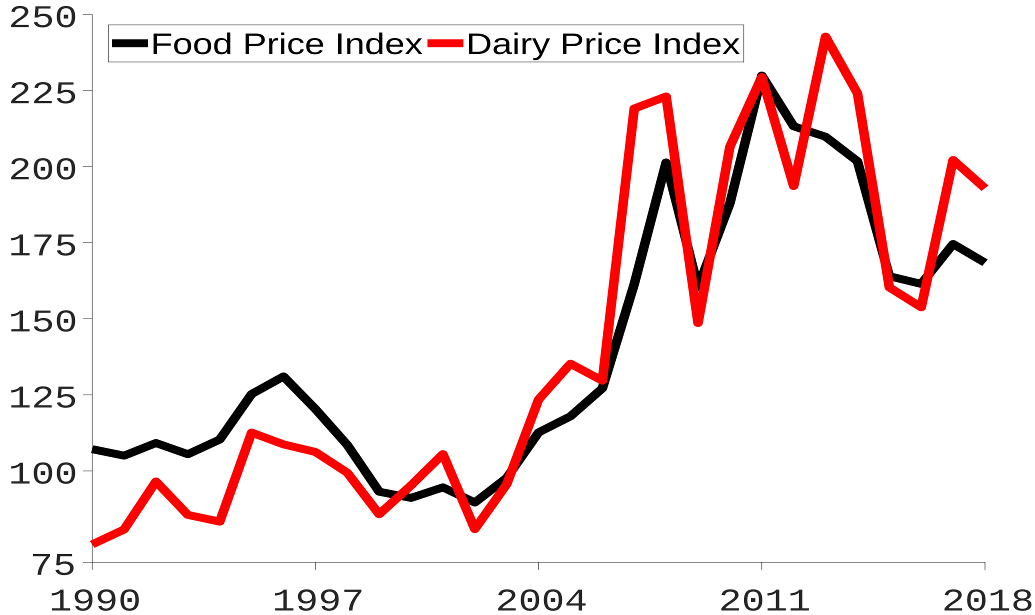

Figure A4: Annual food price indices (2002-2004=100), 1990-2018
Source: `http://www.fao.org/worldfoodsituation/foodpricesindex/en/`