Japan & The World Economy 67 (2023) 101204

Contents lists available at ScienceDirect

Japan & The World Economy

journal homepage: www.elsevier.com/locate/jwe

Evaluation of fiscal policy using alternative GDP data in Japan

Shin-ichi Fukuda

Faculty of Economics, University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

JEL Classification:
E01
E62
H30
E3

Keywords:
GDP statistics
Fiscal policy
Structural changes
Ultra-low interest rates

ABSTRACT

In macroeconomics, the effectiveness of fiscal policy is an old, but new controversial question that has been debated for many years. The controversy naturally arises if analytical frameworks, sample periods, or targeted countries are different. However, it sometimes arises, even when using similar analytical frameworks and sample periods. The purpose of this study is to explore whether revisions to the GDP data series are a source of the controversy in Japan. Since the mid-1990s, the Japanese economy has remained stuck in a liquidity trap where interest rates have fallen to zero. However, not only has the cumulative fiscal deficit increased to an unprecedented level, but many structural problems have also emerged. Under these circumstances, it is unclear whether the fiscal policy has worked effectively in the Japanese economy. When a reduced-form equation or VAR model was estimated, the empirical findings were quite different depending on GDP benchmark year we used in the analysis. When benchmark year 2011 was used, fiscal expenditure was effective under ultralow interest rates. In contrast, when benchmark year 2015 was used, it was not effective, especially since around 2010, even under ultra-low interest rates. This implies that in Japan, the effectiveness of fiscal policy must be carefully interpreted, noting which GDP benchmark year is used in the analysis.

1. Introduction

In macroeconomics, the effectiveness of fiscal policy is an old but new question that has been debated for many years. In the literature, empirical results have largely been controversial regarding effectiveness. The controversy naturally arises if analytical frameworks, sample periods, or targeted countries are different. However, it sometimes arises, even when using similar analytical frameworks and sample periods. The purpose of this study is to explore whether revisions to the GDP data series are a source of the controversy in the Japanese economy.

Japan's GDP data series has been published based on the System of National Accounts (SNA) since 1966. The estimation methods were revised in accordance with the revisions of the international standard when the benchmark year was changed. Benchmark revisions are desirable to reflect developments in the economic and financial environments by incorporating several large-scale detailed source statistics. However, they are accompanied by a variety of changes in their predecessors' concepts and definitions. As a result, the revised GDP data may have different implications for the effectiveness of fiscal policies even if we use similar analytical frameworks and sample periods.

Table 1 summarizes how the GDP data series has been revised over the past few decades in Japan. The latest series from 1955 to 1979 was released on June 21, 2001, and has not been revised thereafter. The series from 1980 to 1993 was revised when the benchmark year changed. However, only simple retroactive time series have been complied since the data series based on benchmark year 2005 was released. The latest main series from 1980 to 1993 was thus released on December 9, 2005. The series from 1994 was revised when the benchmark year changed. The main series from 1994 was automatically replaced with the new series when the data series based on a new benchmark year was released.

Table 2 reports chronological changes in the main series of Japan's GDP data. It suggests that a substantial part of the main series had been replaced with the new series when the benchmark year was changed. For example, until August 29th, 2002, the main series had been based on benchmark year 1990. But on August 30th, 2002, those from 1980Q1 to 2001Q1 were replaced with the series based on benchmark year 1995.[1] Similarly, until December 7th, 2020, the main series was the series based on benchmark year 2011. But on December 8th, 2020, those from 1994Q1 were replaced with the series based on benchmark year 2015. In the former case, the main series for nearly 21 years was replaced with the new series when the benchmark year was changed from 1990 to

E-mail address: sfukuda@e.u-tokyo.ac.jp.

[1] The benchmark year was revised from 1990 to 1995 in October 2000. However, on the Cabinet Office website, the data was updated from the benchmark year 1990–1995 on August 30th, 2002 following the QE reform.

https://doi.org/10.1016/j.japwor.2023.101204
Received 21 May 2022; Received in revised form 30 March 2023; Accepted 22 June 2023
Available online 27 June 2023

Table 1
Benchmark revisions of Japan's GDP.

SNA	2008SNA	2008SNA	1993SNA	1993SNA	1993SNA	1993SNA	1993SNA	1968SNA
benchmark year	**2015**	**2011**	**2005**	**2000**	**2000**	**1995**	**1995**	**1990**
deflators	chain-linking method	chain-linking method	chain-linking method	chain-linking method	fixed-base year method	chain-linking method	fixed-base year method	fixed-base year method
released date	8-Dec-2020	8-Dec-2016	9-Dec-2011	9-Dec-2005	9-Dec-2005	8-Dec-2004	30-Aug-2002	21-Jun-2001
1955–1979	NA	NA	NA	NA	NA	NA	NA	Main Time Series
1980–1993	Retroactive Time Series	Retroactive Time Series	Retroactive Time Series	Main Time Series	Reference Time Series	Retroactive Time Series	Reference Time Series	Reference Time Series
from 1994	Main Time Series	Reference Time Series	Reference Time Series	Reference Time Series	Reference Time Series	Reference Time Series	Reference Time Series	Reference Time Series
the latest data		2020Q3	2016Q3	2011Q3	2006Q1	2005Q3	2005Q2	2001Q1

Table 2
Chronological changes of the main time series.

	1955–1979	1980–1993	from 1994
until 29-Aug-2002	68SNA, 1990, fixed	68SNA, 1990, fixed	68SNA, 1990, fixed
30-Aug-2002–7-Dec-2004	68SNA, 1990, fixed	93SNA, 1995, fixed	93SNA, 1995, fixed
8-Dec-2004–8-Dec-2005	68SNA, 1990, fixed	93SNA, 1995, chain	93SNA, 1995, chain
9-Dec-2005–8-Dec-2011	68SNA, 1990, fixed	93SNA, 2000, chain	93SNA, 2000, chain
9-Dec-2011–7-Dec-2016	68SNA, 1990, fixed	93SNA, 2000, chain	93SNA, 2005, chain
8-Dec-2016–7-Dec-2020	68SNA, 1990, fixed	93SNA, 2000, chain	08SNA, 2011, chain
after 8-Dec-2020	68SNA, 1990, fixed	93SNA, 2000, chain	08SNA, 2015, chain

1995. In the latter case, the main series for nearly 27 years was replaced when the benchmark year was changed from 2015 to 2020. In the following analysis, we explore how the substantial revisions affected the evaluation of the effectiveness of fiscal policy in Japan, paying special attention to the effects of the benchmark year revision on December 8th, 2020.

Considering the effectiveness of macro-fiscal policy, the Japanese economy has exhibited several notable features in recent years that might have affected the effectiveness of fiscal policy. One feature is that since the mid-1990s, the economy has been under a "liquidity trap," in which interest rates have fallen to zero. If fiscal spending raised interest rates, a "crowding out" of private investment would have reduced the effectiveness of the fiscal policy on national income. Thus, fiscal policy becomes relatively effective under the liquidity trap, under which fiscal spending does not raise interest rates. In addition, as Blanchard (2019) indicates, expanding fiscal spending becomes less costly under the liquidity trap because it has limited impacts on fiscal deficits.

However, the accumulation of fiscal deficits in Japan has reached a globally unprecedented level, which other major countries have never experienced. Therefore, although interest rates are historically low, their burden on future generations has become extremely heavy. To the extent that the government is responsible for repaying the deficit, the accumulation of fiscal deficits may reduce the propensity to consume and offset the effect of fiscal policy on national income when people have serious concerns about future burdens.

Furthermore, following the bursting of bubbles, various structural problems have emerged in Japan. For example, the bubble burst in the early 1990s led to the emergence of "zombie companies" that would otherwise have to exit the market.[2] In environments in which persistent financial support for zombie companies delays economic recovery, the expansion of fiscal spending would be less effective. In addition, social security-related expenditures have increased owing to the declining birth rate and an aging population. Thus, a substantial portion of government expenditure is on income transfer, which has been less effective in increasing national income.

In previous studies, Miyamoto et al. (2018) showed that the effectiveness of Japan's fiscal policy greatly increased under a "liquidity trap, " even though Kato et al. (2018) pointed out that tax cuts were less effective. Iwata (2011) showed that the debt-stabilizing tax policies employed in Japan during the 1980s and 1990s played a role in making the short-run multipliers large. Using a multivariate autoregressive (VAR) model, Bayoumi (2001), Kuttner and Posen (2001, 2002), and Morita (2015) found that fiscal policy had significant positive time-varying effects on GDP after adoption of unconventional monetary policy. Using panel data by prefecture, Brückner and Tuladhar (2014) found that the average local government spending multiplier was positive and significantly different from zero. Kameda et al. (2021) estimated the local fiscal multiplier on output to be 1.7 at the regional level.

By contrast, focusing on the structural problem of a rapidly aging society with a declining birthrate, Yoshino and Miyamoto (2017) showed calibration results in which the effectiveness of Japan's fiscal policy declined substantially. Otsu and Shibayama (2022) demonstrated that government spending on aging may have reduced productive resources in Japan. Bessho (2021) provided evidence that local fiscal multiplier was larger in non-aged areas than aged areas. Futagami and Konishi (2018) quantitatively proposed that the debt and deficit reduction rules based only on government consumption and investment expenditure cuts improve households' welfare in the Japanese economy. Werner (2004) and Auerbach and Gorodnichenko (2017) indicated that the multiplier effect of fiscal policy, which was effective during previous decades, has been declining. Fukuda and Yamada (2011) pointed out that the effect of fiscal expansion on stock prices declined as fiscal deficits were accumulated. Using a threshold VAR, Kameda (2014) showed that the effects of government expenditures have diminished since around the 1990s Umeda et al. (2018) found that many Japanese economists believe that the multiplier effect had declined in recent years. Saxegaard et al. (2022) discussed that increased policy uncertainty might have affected macroeconomic performance.

Thus far, empirical analysis has largely been controversial regarding the effectiveness of fiscal policy in Japan. Unfortunately, we cannot

[2] For the impact of zombie companies on the Japanese economy, see Caballero et al. (2008) and Fukuda and Nakamura (2011).

simply compare the results of the above-mentioned studies because they used different analytical frameworks and different sample periods. However, in this study, we update the data of these previous studies and evaluate the effectiveness of fiscal policy estimating a reduced-form equation and VAR model. We show that a source of controversy in the Japanese economy is revisions to the GDP data series.

2. "Government expenditures" used in the analysis

2.1. Concept of "government expenditures"

Before implementing the empirical estimations, this section overviews the concept of "government expenditures" used in the following analysis. "Government expenditures" include "government consumption" and "public investment" based on the "National Accounts (GDP statistics)". They are substantially different from the total expenditures of the central and local governments in two respects.

First, they are expenditures by the "general government," composed not only of central and local governments' expenditures but also of social security funds. The central government's expenditures include those of its general account, special account, and independent administrative agencies. Local government expenditures include those of the government's general account, local public businesses, and local independent administrative agencies. The social security funds include the central government's special account for public pension and employment insurance; the local government's account for medical and nursing care services; a part of the civil servant mutual aid association account; and the Government Pension Investment Fund.

Using the GDP data based on benchmark year 2015 [2008 SNA], Fig. 1 shows government consumption from FY 2005 to FY 2020 classified by central and local governments and social security funds. The central government's share was small, ranging from 14% to 16% throughout the period. In contrast, the share of local governments exceeded 40%. In recent years, however, the social security fund has gradually increased its share from 36% in 2005 to 42% in 2020. Consequently, since FY 2017, the share of social security funds has slightly exceeded that of the local governments.

Second, because GDP is the sum of added value, only those that generate added value are included in "government expenditures." For this reason, land purchases are not included in "public investment," despite being included in the general accounts of central and local governments. Additionally, income transfers from the government to the private sector are not included in government consumption unless they generate added value. Recently, a variety of income transfers have occurred from central and local governments to the private sector. Of these, goods and services purchased by the general government for providing in-kind transfers and those supplied by individual government agencies are included in "government consumption." However, those provided as monetary transfers are not included in "government consumption."

2.2. Changes in "government expenditures"

Fig. 2 shows how the ratios of "government consumption" and "public investment" to GDP have evolved since FY 1980, using real and seasonally adjusted quarterly data. Public investment, which was around 10% of GDP until the end of the 1990s, significantly decreased in the early 2000s, reflecting reductions in public works spending; it has been around 5% of GDP since FY 2007. However, government consumption has been steadily increasing since the 1990s, from approximately 15% of GDP in the early 1990s to around 20% since 2009.

Fig. 3 depicts the changes in "government consumption" in nominal terms by function since FY 2005. This shows that expenditures for general administration have been greatly reduced in recent years, from nearly 11 trillion yen in FY 2005 to around 9 trillion yen in FY 2012 and thereafter. However, there has been a huge increase in "health" expenditures, such as medical expenses and "social assistance" expenditures, such as welfare benefits. "Health" expenditures were about 30 trillion yen in FY 2005, but they exceeded 41 trillion yen in FY 2019 and 2020. "Social assistance" expenditures were about 9.4 trillion yen in FY 2005 but exceeded 15 trillion yen in FY 2020. The increase in social security-related expenditures was a major factor in the recent increases

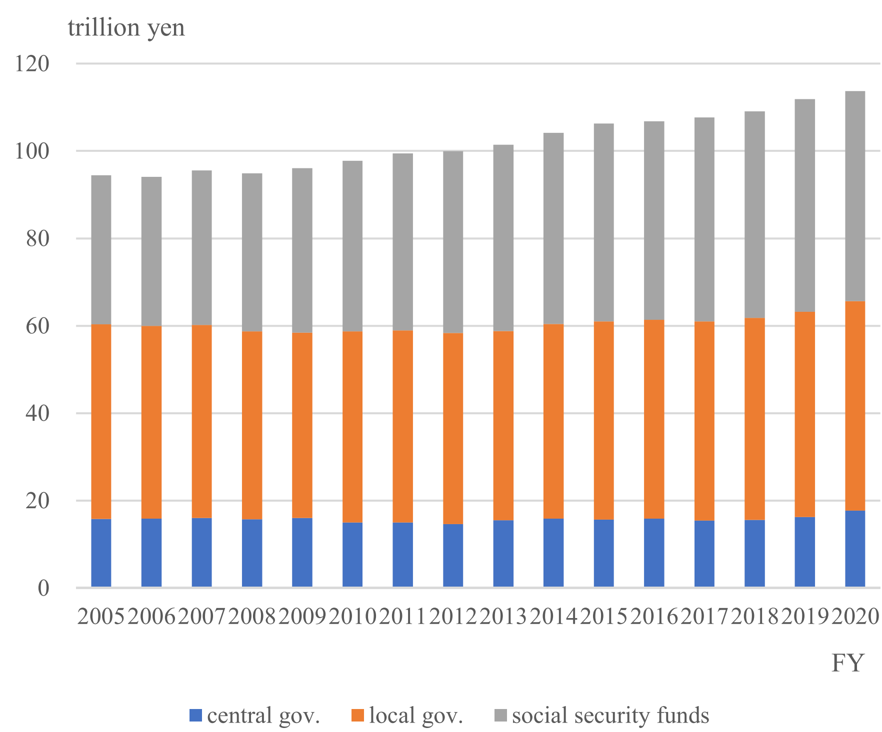

Fig. 1. Components of government consumption.
Source: Annual Report on National Accounts, Cabinet Office.

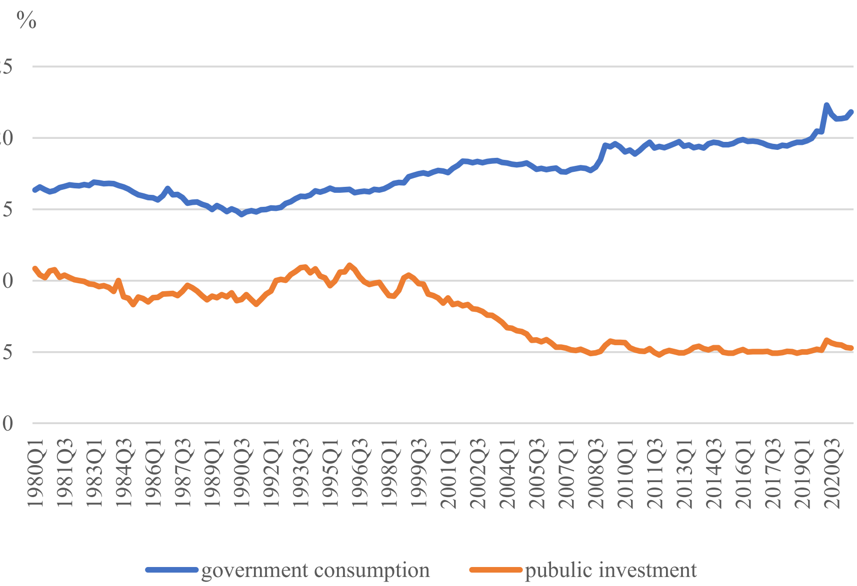

Fig. 2. The ratios of government consumption and public investment to GDP.
Source: Annual Report on National Accounts, Cabinet Office.

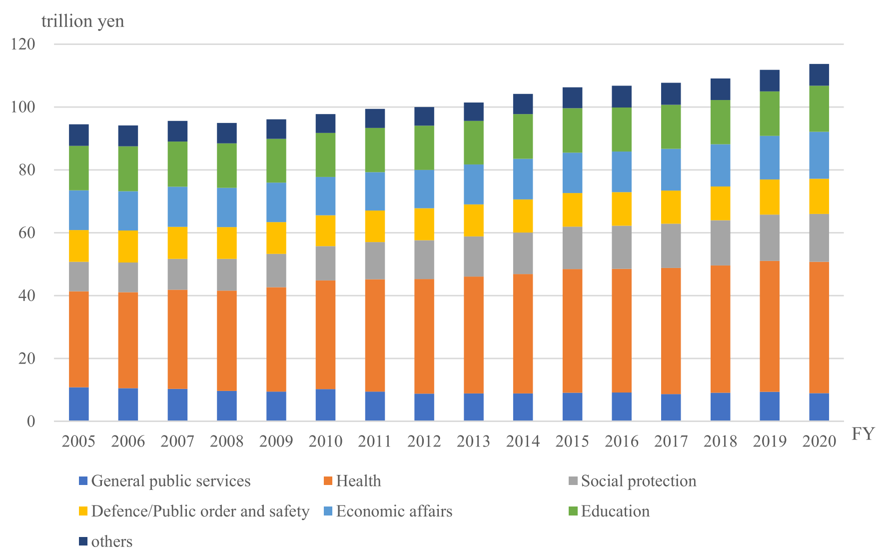

Fig. 3. Government consumption classified by function.
Source: Annual Report on National Accounts, Cabinet Office.

in "government consumption."

Fig. 4 shows how the short-term fluctuations in "government consumption" and "public investment" have changed since FY 1980 using the quarterly (annualized) growth rates of their real and seasonally adjusted data. From the figure, we can see that public investment has been much more volatile than government consumption, especially in the 1980s and 1990s. In the 2000s, short-term fluctuations significantly decreased in both "government consumption" and "public investment." After the collapse of Lehman Brothers in 2008 and the Great East Japan Earthquake in 2011, short-term fluctuations of public investment temporarily increased. However, the trend decline of short-term fluctuations continued throughout the 2000s.

3. Correlations of GDP data series based on different benchmark years

The purpose of this study is to investigate whether revisions to the GDP data series change the evaluation of fiscal policy in Japan. As explained in the Introduction, benchmark revisions to the GDP data series were accompanied by a variety of changes in the concepts and definitions. As a result, the revised GDP data may have different features even when we use the sample periods. This section explores whether the revised GDP data series has different features by calculating the correlations. The correlation coefficient is calculated using the growth rates of seasonally adjusted quarterly data on real GDP and its components. The real values were calculated using both the fixed-base year method and the chain-linking method for benchmark years 1995 and 2000. However, they were calculated only by the fixed-base year method for benchmark year 1990, and only by the chain-linking method for benchmark years 2005, 2011, and 2015. Although they were constructed using different methods and benchmark years, they commonly provide time-series data from the first quarter of 1980, including the simple retroactive time series, and the first quarter of 1994, including only the main series.

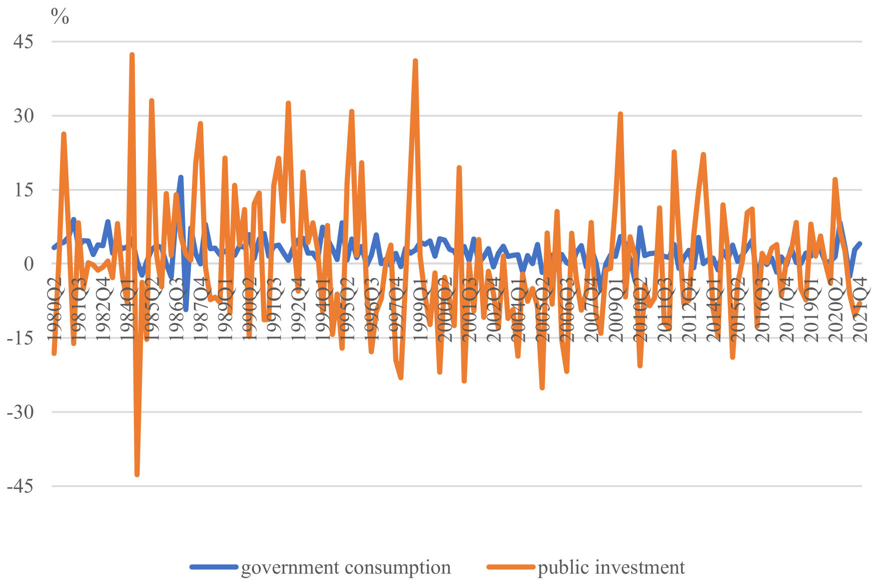

Fig. 4. The growth rates of government consumption and public investment.
Source: Quarterly Estimates of GDP (benchmark year 2015 (2008SNA)), Cabinet Office.

Table 3
Correlation coefficient of GDP statistics by different benchmark years.

Benchmark year	Sample period	GDP	Priv. con.	Priv. Inv.	Res. Inv.	Gov. con.	Gov. inv.
2011, chain	1980Q2–2020Q3	0.992	0.988	0.969	0.977	0.981	0.964
2005, chain	1980Q2–2016Q3	0.940	0.915	0.829	0.973	0.910	0.891
2000, chain	1980Q2–2011Q3	0.926	0.897	0.839	0.954	0.855	0.848
2000, fixed	1980Q2–2006Q1	0.853	0.887	0.808	0.960	0.881	0.791
1995, chain	1980Q2–2005Q3	0.779	0.805	0.703	0.960	0.499	0.535
1995, fixed	1980Q2–2005Q2	0.811	0.859	0.822	0.957	0.734	0.538
1990, fixed	1980Q2–2001Q1	0.541	0.657	0.662	0.837	0.625	0.538
Benchmark year	Sample period	GDP	Priv. con.	Priv. Inv.	Res. Inv.	Gov. con.	Gov. inv.
2011, chain	1994Q1–2020Q3	0.992	0.990	0.945	0.978	0.959	0.938
2005, chain	1994Q1–2016Q3	0.931	0.865	0.736	0.981	0.847	0.826
2000, chain	1994Q1–2011Q3	0.907	0.788	0.760	0.960	0.699	0.791
2000, fixed	1994Q1–2006Q1	0.805	0.792	0.702	0.977	0.742	0.654
1995, chain	1994Q1–2005Q3	0.758	0.761	0.686	0.979	0.638	0.576
1995, fixed	1994Q1–2005Q2	0.774	0.749	0.710	0.978	0.728	0.578
1990, fixed	1994Q1–2001Q1	0.355	0.629	0.523	0.849	0.369	0.678

Note. The correlation coefficients were calculated using their growth rates.

Table 3 reports how the growth rate of real GDP and its components constructed by benchmark year 2015 were correlated with those constructed by previous five benchmark years: 1990, 1995, 2000, 2005, and 2011. The correlations are calculated for the growth rates (logged differences) of each variable from the second quarter of 1980 in Table 3-(1) and for those from the first quarter of 1994 in Table 3-(2). Real GDP and its components based on benchmark year 2015 were highly correlated with those based on benchmark year 2011. However, they did not have high correlations with those of other benchmark years. In particular, they were less correlated with those based on benchmark year 1990, especially for those from the first quarter of 1994. Among the components of real GDP, residential investment had relatively high correlations. In contrast, government consumption and public investment had relatively low correlations. In the case of real GDP and government consumption, the correlations between benchmark years 2015 and 1990 were only 0.355 and 0.369, respectively, when using the data from the first quarter of 1994.

Benchmark revisions frequently changed the definition of GDP substantially. For example, the 1995 revision newly included order-made software and consumption of fixed capital for social capital, while the 2000 revision included packaged software. The 2005 revision included self-developed software and financial intermediary services (FISIM). The 2011 revision included investments in research and development, patents and other services investments, and changes in the treatment of defense equipment. Similarly, benchmark revisions changed the definition of government consumption and public investment substantially. For example, 1995 benchmark revision newly included the transfer of social insurance benefits for medical and long-term care expenses in government consumption. The following revisions also changed the accounting of the consumption of fixed capital of social capital in government consumption. The 2011 benchmark revision newly recorded R&D investment in public investment. The 2015 benchmark revision reflected retroactive revisions of the public investment. It is likely that these revisions made the correlations low between benchmark years 2015 and 1990.

As shown in Table 2, the main series from 1955 to 1979 are based on benchmark year 1990, those from 1980 to 1993 are based on benchmark year 2000 (chain-linking method), and those from 1994 are based on benchmark year 2015. In the literature, when analyzing policy effects from a medium- to long-term perspective, many previous studies have connected these data series, each of which was constructed using different methods. However, the above low correlations suggest that it is

Table 4
Correlation coefficient between GDP and government expenditures.

(1) 1980Q2-2001Q1								
	15, chain	15, chain	11, chain	00, chain	00, fixed	95, chain	95, fixed	90, fixed
Gov. con.	0.206	0.216	0.180	0.161	0.147	-0.014	-0.004	-0.056
Gov. inv.	0.196	0.165	0.108	0.082	0.084	0.145	0.097	0.297
Gov. exp.	0.254	0.232	0.161	0.134	0.128	0.130	0.091	0.267

(2) 1980Q2-2016Q3			
	15, chain	11, chain	05, chain
Gov. con.	0.249	0.257	0.208
Gov. inv.	0.169	0.183	0.149
Gov. exp.	0.246	0.256	0.194

(3) 1994Q1-2016Q3			
	15, chain	11, chain	05, chain
Gov. con.	0.224	0.229	0.099
Gov. inv.	0.057	0.110	0.126
Gov. exp.	0.122	0.150	0.117

Note. The correlation coefficients were calculated using their growth rates.

not appropriate to connect them to construct a medium- to long-term GDP data series. Connecting the GDP statistics constructed using different methods may likely cause serious problems.

Table 4 shows the correlation coefficients between real GDP and government expenditure (i.e., government consumption, public investment, and their sum) based on six alternative benchmark years. When using the data from the second quarter of 1980 to the first quarter of 2001, the correlation coefficients were relatively high based on benchmark years 2011 and 2015, and relatively low based on benchmark year 1995. Even when using data until the third quarter of 2016, the correlation coefficients varied depending on benchmark year. The correlation coefficients are relatively high based on benchmark year 2011 and relatively low based on benchmark year 2005. They suggest that revised GDP data may have different implications for the role of fiscal policy, even if we use the same sample period.

4. Structural breaks

In the previous section, we demonstrated that the real GDP and its components based on benchmark year 2015 had relatively high correlations with those based on benchmark year 2011. However, it is unclear whether these high correlations were stable throughout the sample period. This section examines whether there are any structural changes in these correlations. When examining structural changes during the sample period, it is important to determine when they have occurred. We use the Quandt–Andrews structural break test (Andrews, 1993), typically used when the timing of a structural change is unknown, to determine when a structural change has occurred.

In the analysis, we regressed the growth rate of each variable based on benchmark year 2015 and benchmark year 2011 for two alternative sample periods: the second quarter of 1980 to the fourth quarter of 2019 and the first quarter of 1994 to the fourth quarter of 2019. We then applied the Quandt–Andrews test to identify when a structural change occurred in each regression for each sample period.

$$\Delta logX(2015) = constant + a\ \Delta logX(2011), \quad (1)$$

where $X(2011)$ is the real GDP or its components based on benchmark year 2011, and $X(2015)$ is based on benchmark year 2015.

Table 5 summarizes the results of the structural beak test for the two alternative sample periods: from the second quarter of 1980 to the fourth quarter of 2019 and from the second quarter of 1980 to the fourth quarter of 2019. It reports the identified structural breakpoint in each regression and its maximum Wald F-statistics. In both sample periods, the maximum Wald F-statistic indicated that the correlations had no significant structural change when we used real GDP, private consumption, private investment, and residential investment. However, for the sample period from the second quarter of 1980 to the fourth quarter of 2019, this indicates that there was a significant structural change when we used government consumption, public investment, and their sum. The timing of the structural change was the fourth quarter of 2013 for government consumption, the fourth quarter of 2011 for public investment, and the third quarter of 2009 for government expenditure (i.e., government consumption + public investment).

The results of the structural beak test are essentially the same for the sample period from the first quarter of 1994 to the fourth quarter of 2019: Because of the relatively small sample size, the maximum Wald F-statistic was relatively small in each regression. The correlations had no significant structural change, not only for real GDP, private consumption, private investment, and residential investment but also for government consumption. However, the identified structural changes were still significant for both public investment and government expenditure. More importantly, the identified structural change points are the same as those from the second quarter of 1980 to the fourth quarter of 2019 for government consumption, public investment, and government expenditure.

Table 6 reports the correlation coefficients between real GDP and government expenditures (that is, government consumption, public investment, and their sum) based on two benchmark years, 2011 and 2015. For each government expenditure, it shows the correlation coefficients for the period before and after the structural break. When using the data before the structural break, the two benchmark years showed similar correlation coefficients. However, when using the data

Table 5
Quandt–Andrews structural break test.

	GDP	Priv. con.	Priv. Inv.	Res. Inv.	Gov. con.	Gov. inv.	Gov. exp.
1980Q2–2020Q3	1988Q3	2013Q1	1992Q3	2014Q3	2013Q4	2011Q4	2009Q3
	(5.10)	(6.71)	(2.64)	(5.96)	(18.30)***	(28.19)***	(15.96)***
1994Q1–2020Q3	2008Q3	2014Q3	2011Q3	2014Q3	2013Q4	2011Q4	2009Q3
	(1.78)	(2.43)	(0.84)	(6.25)	(7.64)	(16.07)***	(10.19)*

Note) Maximum Wald F-statistics are in parentheses. *= 10%, ** = 5%, ***= 1%.

Table 6
Structural break and correlation between GDP and government expenditures.

		15, chain	11, chain
Gov. con.	1980Q2–2013Q4	0.227	0.236
	1994Q1–2013Q4	0.189	0.199
	2014Q1–2020Q3	0.391	0.497
Gov. inv.	1980Q2–2011Q4	0.160	0.146
	1994Q1–2011Q4	0.003	0.011
	2012Q1–2020Q3	-0.056	0.167
Gov. exp.	1980Q2–2009Q3	0.253	0.233
	1994Q1–2009Q3	0.109	0.097
	2009Q4–2020Q3	0.109	0.363

after the structural break, benchmark year 2011 showed higher correlation coefficients than benchmark year 2015. It suggests that the revised GDP data may have different implications for the role of fiscal policy even if we use the same sample period. The effectiveness of fiscal policy should be carefully interpreted noting which benchmark year is used in the analysis.

The period after the structural break corresponds to that in which long-term interest rates had a sharp downward trend toward zero in Japan. Thus, it is important to see whether the fiscal policy was effective during this period. However, since the data based on different benchmark years cause different time-series properties of government expenditures, we likely have different implications for the effectiveness of fiscal policy depending on benchmark year we use in the analysis. Structural changes in the correlations of government expenditure raise serious concerns in empirical studies in Japan.

In the GDP statistics, benchmark years 2011 and 2015 rely on the 2008 SNA for the System of National Accounts. However, significant revisions have been made, such as the inclusion of renovations in investment, the reflection of sales margins for condominiums, and the updating of benchmarks for construction output estimates. More importantly, as we discuss in Section 9, 2015 benchmark revision reflected retroactive revisions of the public construction data since 2011. Thus, not only did the level of GDP rise by about 1.3%, but the short-term fluctuations in government expenditures also changed significantly. The contradictory results might have occurred because of these changes.

5. The estimation of reduced-form equations

5.1. Analytical framework

In the previous sections, we examined the correlations of alternative GDP data and suggested that the use of different benchmark years may have different implications for the Japanese economy. In this section, we estimate a reduced-form equation and examine whether revisions to the GDP data series can have different implications for the effectiveness of fiscal policy in Japan.

The estimation of a reduced-form equation is limited in that its economic structure is represented by a black box. In addition, the estimated “effect” only indicates a “Granger causality,” which does not necessarily imply a true causal relationship. However, the estimation of a reduced-form equation is the simplest method to determine the effect of exogenous variables and has been extensively used in previous research. Using seasonally adjusted quarterly data, we estimate the following equation, including the lagged dependent variables:

$$\Delta Y_t = \text{constant} + \sum_{i=0}^{1} a_i \Delta G_{t-i} + \sum_{i=1}^{2} b_i \Delta Y_{t-i} \quad (2)$$

Here, ΔY_t is the growth rate (logged difference) of real GDP at time t and ΔG_t is the growth rate of real government expenditure at time t.

We estimate Eq. (2), including the control and dummy variables. The control variables are the growth rates of real exports and real money balances (*M2* divided by the consumer price index). We use current and lagged values for real exports and one and two lagged values for real money balances. Dummy variables are included to remove the effects of real GDP outliers. Specifically, we include four dummy variables, each of which takes one in the first quarter of 2009, the first quarter of 2011, the second quarter of 2014, and the fourth quarter of 2019, and zero otherwise. The first quarter of 2009 and the first quarter of 2011 reflected the output decline due to the Global Financial Crisis and the Great East Japan Earthquake, respectively. The second quarter of 2014 and the fourth quarter of 2019 reflect the decline in output following the consumption tax hikes.

In the analysis, we used seasonally adjusted GDP data series based on benchmark years 2011 and 2015, from the first quarter of 1994 to the fourth quarter of 2019. The analysis did not include data from 2020 to exclude the enormous impact of COVID-19 on the GDP. Using the two subsample periods identified by the Quandt–Andrews test in the previous section, we explore how the effects of fiscal spending changed in the two subsample periods. Estimation was performed using the ordinary least-squares method.

5.2. Estimation results

Table 7 summarizes the estimation results based on the two alternative benchmark years for the two subsample periods: the first quarter of 1994 to the third quarter of 2009, and the fourth quarter of 2009 to the fourth quarter of 2019. For the first subsample period, the data based on the two benchmark years yielded similar estimation results. In both estimates, the growth rates of real government expenditure and real exports had significant instantaneous impacts, while the two-lagged dependent variable had a significant impact. The estimated coefficients are also similar, except for lagged real exports. In particular, the instantaneous and lagged impacts of real government expenditure on real GDP are very close between the two estimates. This implies that benchmark year revisions to the GDP data series did not change the essential features of the estimation results, especially the effect of fiscal policy, before the structural break.

By contrast, for the second subsample period, the data based on the two benchmark years led to different estimation results. In both estimates, the growth rates of real exports and real money balances have similar significant impacts. However, the growth rates of real government expenditure and the lagged dependent variable showed different impacts in terms of magnitude and statistical significance. Most notably, the impact of fiscal policy showed substantial differences between the two estimates. When comparing the instantaneous impact of real government expenditures, the data based on benchmark year 2011 showed a large and significant impact, but those based on benchmark year 2015 did not. The data based on benchmark year 2011 showed more than 70% larger instantaneous impacts than those based on benchmark year 2015. The benchmark year revision to the GDP data series led to different implications for the effectiveness of fiscal policy after the structural break.

6. Estimations by VAR

In the previous section, we estimated a reduced-form equation and showed that the revision to the GDP data series had quite different implications for the effectiveness of fiscal policy in Japan. In this section, we estimate a vector autoregressive (VAR) model to examine this issue. As with the reduced-form equation, the economic structures behind a VAR model are in a black box, and the “effect” shown by the VAR model only indicates a “Granger causality”. Whereas a reduced-form equation examines the unidirectional “impact” of government spending on GDP, a VAR examines the overall “impacts” of government spending on GDP through various macro variables.

In the following VAR model, we use ΔY_t ($\equiv$ the growth rate of real GDP at time t), ΔG_t ($\equiv$ the growth rate of real government expenditure at time t), and $\Delta(M2/P)_t$ ($\equiv$ the growth rate of real money balances at

Table 7
The effects of fiscal spending on GDP.

	Sample period: 1994Q4 - 2009Q3				Sample period: 2009Q4 - 2019Q4			
	2011, Chain		2015, Chain		2011, Chain		2015, Chain	
	coeff.	t value	coeff.	t value	coeff.	t value	coeff.	t value
C	0.001	0.435	0.001	0.392	-0.001	-0.720	-0.002	-0.822
ΔG	0.170	2.283**	0.179	2.448**	0.375	2.592**	0.216	1.548
ΔG(−1)	0.027	0.359	0.022	0.295	-0.131	-1.009	-0.022	-0.161
ΔY(−1)	-0.106	-0.767	-0.064	-0.467	-0.060	-0.422	-0.144	-0.925
ΔY(−2)	0.160	1.774*	0.153	1.691*	-0.262	-2.201**	-0.123	-1.037
ΔEX	0.156	5.758***	0.157	5.709***	0.176	6.980***	0.179	6.459***
ΔEX(−1)	0.001	0.041	-0.001	-0.021	0.016	0.541	0.008	0.242
Δ(M2/P)(−1)	0.051	0.373	0.056	0.400	0.291	1.840	0.381	2.239**
Δ(M2/P)(−2)	-0.142	-1.060	-0.169	-1.234	0.281	1.684	0.204	1.146
Dum09Q1	-0.007	-0.652	-0.007	-0.599				
Dum11Q1					-0.011	-1.954*	-0.010	-1.720*
Dum14Q2					-0.019	-3.377***	-0.020	-3.556***
Dum19Q4					-0.018	-3.622***	-0.023	-4.322**
Adj. R^2	0.624		0.615		0.727			0.697

Note) ***= significant at the 1% level, ** = significant at the 5% level, and *= significant at the 10% level.

time *t*) as endogenous variables, and ΔEX_t (≡ the growth rate of real exports at time *t*) and four dummy variables as exogenous variables.[3] Specifically, we estimate the following second-order VAR model.

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + B\, Z_t, \qquad (3)$$

Here, $X_t \equiv \{\Delta G_t, \Delta Y_t, \Delta (M2/P)_t\}$ and $Z_t \equiv$ {constant term, ΔEX_t, the four outlier dummies}.

As in the previous section, we estimated the model using seasonally adjusted quarterly GDP data based on benchmark years 2011 and 2015 for two subsample periods: the first quarter of 1994 to the third quarter of 2009 and the fourth quarter of 2009 to the fourth quarter of 2019. Identification was performed by Cholesky decomposition, where the order of exogeneity was ΔG_t, ΔY_t, and $\Delta (M2/P)_t$.

In Fig. 5, solid lines in 5–1 show the cumulative impulse response functions of ΔY_t to ΔG_t for the first subsample period based on the two benchmark years. The data based on the two benchmark years yielded similar results. In both data series, the estimated cumulative impulse response function temporarily approached 0.003 in the second quarter after the shock and was stable around 0.002 thereafter. This implies that both data series show persistent impacts of government spending on GDP in either estimate. The estimated cumulative impulse response function was statistically significant only in the first and second quarters in both the data series.

Solid lines in 5–2 show the cumulative impulse response functions of ΔY_t to ΔG_t for the second subsample period. Unlike the first subsample period, the data based on the two benchmark years led to different results for the second subsample period. The data based on benchmark year 2011 showed significant impacts of government spending on the GDP. That is, the estimated cumulative impulse response function temporarily approached 0.003 in the first quarter after the shock and was stable at approximately 0.002 thereafter. More notably, it was statistically significant not only in the first quarter, but also at five and six quarters after the shock. In contrast, data based on benchmark year of 2015 did not show a significant impact of government spending on GDP. In other words, the estimated cumulative impulse response function was always less than 0.002 and was stable at approximately 0.0012 after the sixth quarter. Notably, it was not statistically significant for any quarter after the shock. The data based on benchmark year 2011 showed a nearly 70% larger cumulative impulse response in the tenth quarter than those based on benchmark year 2015. Similar to the reduced-form equation, the VAR model confirmed that the revision to the GDP data series changed the evaluation of fiscal policy after the structural break.

7. The effects of revisions to annual GDP

In the previous sections, we showed that the revisions to quarterly GDP and its components might have different implications for the effectiveness of fiscal policy in Japan. In this section, we examine whether the implications are valid using annual data. Two types of annual data are available for the GDP data series. One is annual data based on the fiscal year, and the other is based on the calendar year. Data based on the fiscal year might be a desirable property to measure government expenditures because the government sectors usually make their decisions based on each fiscal year. However, the data based on calendar years might also be useful because a substantial part of GDP relies on data based on each calendar year. We explore whether revisions to the annual GDP data series might have different implications using both fiscal year and calendar year data.

Table 8 reports how real GDP and its components based on benchmark year 2015 were correlated with those of previous five benchmark years: 1990, 1995, 2000, 2005, and 2011. The correlations are calculated for the growth rate (logged difference) of each variable from the second quarter of 1980. In both fiscal and calendar year data, the annual data based on benchmark year 2015 were highly correlated with those of previous benchmark years, except for government consumption based on benchmark years 1990 and 1995. In most cases, the correlation coefficient of each annual data exceeded 0.9, which was much higher than that of the quarterly data in Table 3. This implies that unlike quarterly GDP and its components, annual GDP and its components were little affected by the revisions.

However, GDP revisions changed the correlations between real GDP and government expenditures, even in the annual data. Using data from 1981 to 2000, In Table 9, the table in 9–1 shows the correlation coefficients between real GDP and government expenditures (i.e., government consumption, public investment, and their sum) for six benchmark years. The correlation coefficients based on benchmark year of 2015 were relatively similar to those based on benchmark years of 2011 and 2015. However, they were different from those based on benchmark years of 1990, 1995, and 2000. The results hold true for both fiscal and calendar year data.

More notably, when considering a structural change, the revision to calendar year GDP changed the correlations between real GDP and government expenditure after 2010. In Table 9, the table in 9–2 shows the correlation coefficients for two benchmark years, 2011 and 2015, for two subsample periods:1994–2009 and 2010–2019. For the first subsample, the correlation coefficients were similar between benchmark

[3] The four dummy variables are those used in the previous section to remove the effects of real GDP outliers.

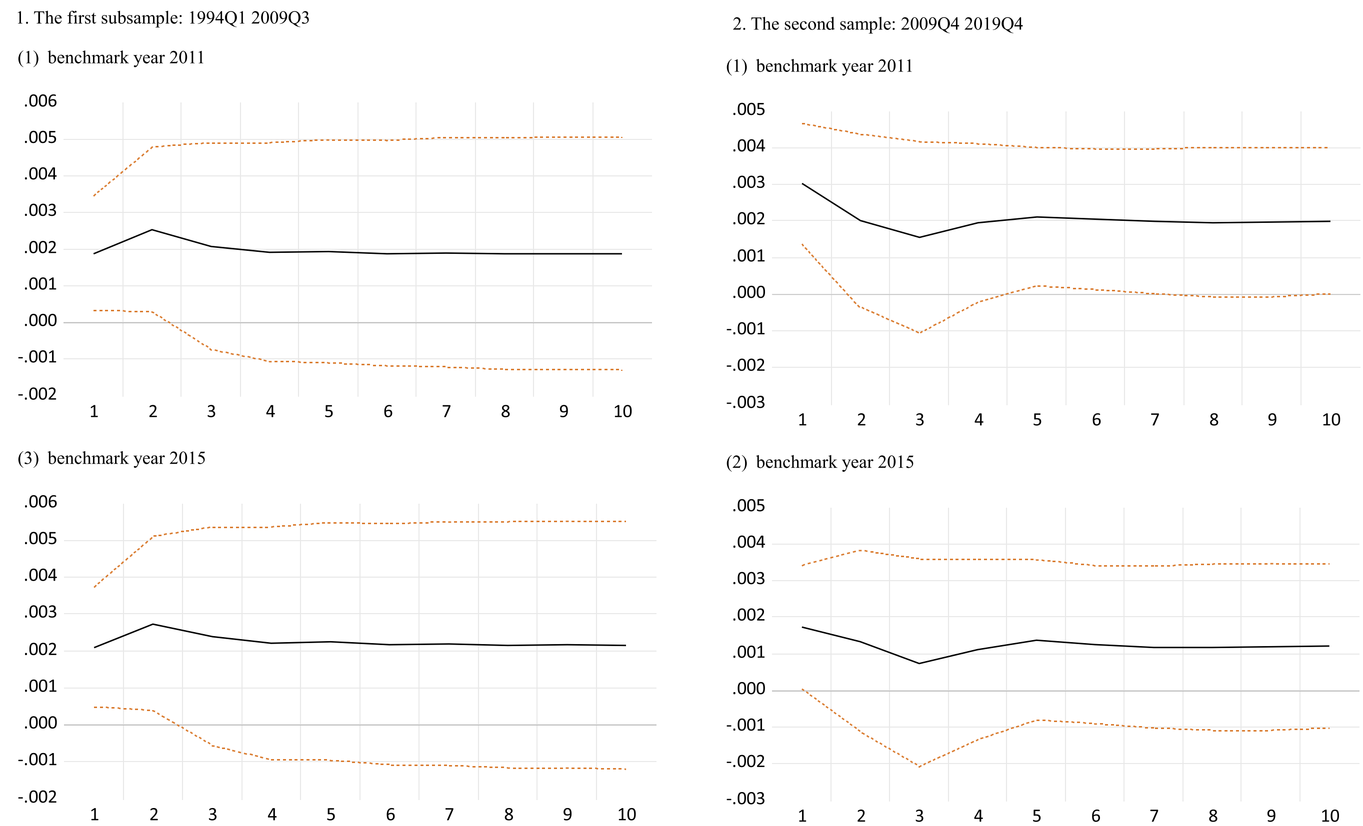

Fig. 5. Cumulative impulse response functions of ΔY_t to ΔG_t. Note: The dotted lines ±2σ denote standard errors.

Table 8
Correlation coefficient of annual GDP statistics by different benchmark years.

(1) fiscal year

		GDP	Priv. Con.	Priv. Inv.	Res. Inv.	Gov. Con.	Gov. Exp.
11, chain	until 2019	0.997	0.998	0.996	0.990	0.980	0.994
05, chain	until 2015	0.983	0.982	0.944	0.985	0.943	0.965
00, chain	until 2010	0.983	0.980	0.950	0.975	0.949	0.958
95, chain	until 2004	0.949	0.952	0.937	0.985	0.819	0.879
90, fixed	until 2000	0.923	0.954	0.926	0.979	0.467	0.915

(2) calendar year

		GDP	Priv. Con.	Priv. Inv.	Res. Inv.	Gov. Con.	Gov. Inv.	Gov. Exp.
11, chain	until 2019	0.996	0.996	0.994	0.984	0.984	0.992	0.994
05, chain	until 2015	0.981	0.975	0.928	0.974	0.941	0.963	0.965
00, chain	until 2010	0.984	0.972	0.934	0.961	0.957	0.948	0.958
95, chain	until 2004	0.969	0.954	0.937	0.971	0.478	0.923	0.879
90, fixed	until 2000	0.917	0.963	0.935	0.968	0.480	0.907	0.915

years of 2011 and 2015 for both the fiscal and calendar year data. Even for the second subsample, they are similar to the fiscal year data. However, when using calendar-year data, they become different for the second subsample. The results imply that a structural change in calendar year GDP changed the correlation between real GDP and government expenditure after 2010.

8. The instrumental variable method

In this section, we examine whether the results in Section 5 are valid even when we estimate Eq. (2) using the instrumental variable method. The instrumental variable method is useful not only when some of the explanatory variables are endogenous but also when some of the explanatory variables include measurement errors. In the GDP data series, different benchmark years may have entailed different time-series properties because of measurement errors. If this is the case, the instrumental variable method would be a powerful tool for fixing estimation biases.

The instrumental variables used in the following analysis are the constant term, four outlier dummy variables, lagged dependent variables (i.e., ΔY_{t-1}, ΔY_{t-2}, ΔY_{t-3}, and ΔY_{t-4}), lagged explanatory variables (i.e., ΔG_{t-1}, ΔG_{t-2}, ΔG_{t-3}, ΔG_{t-4}, ΔG_{t-5}, $\Delta (M2/P)_{t-1}$, $\Delta (M2/P)_{t-2}$, $\Delta (M2/P)_{t-3}$, ΔEX_t, ΔEX_{t-1}, and ΔEX_{t-2}), and lagged real investments (i.e., ΔI_{t-1}, ΔI_{t-2}, and ΔI_{t-3}), where $\Delta I_t \equiv$ the growth rate of real private investment at time t. Table 10 summarizes the estimation results of the instrumental variable method using GDP data based on two alternative benchmark years, 2011 and 2015. Since the two benchmark years showed different results only for the second subsample period in Table 7, we estimate Eq. (2) for the subsample period from the fourth quarter of 2009 to the fourth quarter of 2019.

Table 9
Correlation between annual GDP and government expenditures.

9–1. From 1981–2000

(1) fiscal year

	15, chain	11, chain	05, chain	00, chain	95, chain	90, fixed
Gov. Con.	0.271	0.268	0.253	0.188	-0.036	0.076
Gov. Inv.	0.088	0.070	0.064	-0.135	-0.131	-0.031
Gov. Exp.	0.164	0.133	0.101	0.019	0.046	0.024

(2) calendar year GDP

	15, chain	11, chain	05, chain	00, chain	95, chain	90, fixed
Gov. Con.	0.258	0.253	0.252	0.174	-0.256	0.076
Gov. Inv.	0.052	0.024	0.037	-0.109	-0.126	0.056
Gov. Exp.	0.193	0.159	0.081	-0.018	-0.172	-0.106

9–2. From 1994–2019

(1) fiscal year

	1994–2009		2010–2019	
	15, chain	11, chain	15, chain	11, chain
Gov. Con.	0.369	0.289	0.249	0.281
Gov. Inv.	-0.249	-0.232	-0.059	-0.078
Gov. Exp.	-0.012	-0.017	-0.458	-0.405

(2) calendar year

	1994–2009		2010–2019	
	15, chain	11, chain	15, chain	11, chain
Gov. Con.	0.202	0.156	-0.022	0.112
Gov. Inv.	-0.422	-0.425	-0.028	0.153
Gov. Exp.	-0.220	-0.228	-0.502	-0.377

Table 10
The estimations by instrumental variable method.

	Sample period: 2009Q4 - 2019Q4			
	2011, Chain		2015, Chain	
	coeff.	t value	coeff.	t value
C	-0.002	-1.138	-0.003	-1.302
ΔG	0.549	3.071***	0.427	2.282**
ΔG(−1)	-0.130	-0.979	-0.021	-0.149
ΔY(−1)	-0.033	-0.224	-0.125	-0.773
ΔY(−2)	-0.322	-2.544**	-0.174	-1.375
ΔEX	0.177	6.846***	0.184	6.362***
ΔEX(−1)	0.024	0.762	0.010	0.301
Δ(M2/P)	0.279	1.718*	0.420	2.359**
Δ(M2/P)(−1)	0.341	1.957*	0.245	1.315
Dum11Q1	-0.009	-1.551	-0.008	-1.372
Dum14Q2	-0.016	-2.730**	-0.018	-2.927***
Dum19Q4	-0.018	-3.452***	-0.022	-3.945***
Adj. R^2	0.713		0.673	

Note) ***= significant at the 1% level, ** = significant at the 5% level, and *= significant at the 10% level.

The estimation results in Table 10 show that both the instantaneous and total effects of government expenditures were similar when using the data for benchmark years 2015 and 2011. In other words, the data on benchmark year 2015 similarly detected significantly positive instantaneous and total effects of government expenditure when estimated using the instrumental variable method. The results are in marked contrast with those in Table 7, where both the instantaneous and total effects were lower when using the data of benchmark year 2015 than when using the data of benchmark year 2011 in the estimations. When using the data of benchmark year 2015, the negative correlation coefficient between ΔG_t and the error term or measurement errors in ΔG_t might have caused smaller estimates for both the instantaneous and total effects when estimated by the ordinary least-squares method. These results indicate that the instrumental variable method may fix the estimation biases caused by benchmark year revisions and derive appropriate evaluations of the impacts of fiscal spending in Japan.

9. The benchmark revisions and the effectiveness of fiscal policy

9.1. The discrepancy in government expenditures

Our empirical findings in the previous sections suggest that, in Japan, the effectiveness of fiscal policy has become quite different depending on the GDP benchmark year used in the analysis. When 2011 was used as the benchmark year, fiscal expenditure was effective under ultra-low interest rates. In contrast, when the benchmark year was 2015, it was not effective, especially after 2010, even under ultra-low interest rates. These contrasting results are particularly true when quarterly data are used. This section explores why the estimated effectiveness of fiscal policy differs depending on the GDP benchmark year.

Figs. 6(a)–(b), 7(a)–(b), and 8(a)–(b) depict the quarterly time series of real GDP, real government consumption, and real public investment, respectively, based on benchmark years 2011 and 2015. To allow for a structural break, we depict each series separately for the two subsample periods:1994Q1 to 2009Q4 and 2010Q1 to 2019Q3. The figures show that the GDP and public investment based on 2015 as the benchmark year were always larger than those based on 2011 as the benchmark year. This implies that the benchmark revision persistently increased GDP and public investment levels. In the 2015 revision, investments in the renovation and renewal of buildings were included. This likely contributed to the substantial upward revision of GDP and public investment compared with benchmark year 2011. However, we do not observe such a persistent increase in government consumption. For the first subsample, government consumption shows only a marginal discrepancy in the first half and no conspicuous discrepancy in the second half between the two benchmark years. In contrast, for the second subsample, government consumption based on benchmark year 2015 started to increase faster than that based on benchmark year 2011 around 2013Q4 and remained larger after 2017. This implies that the benchmark revision from 2011 to 2015 brought about a structural change in government consumption around the fourth quarter of 2013 because it caused a structural change in its growth rate. In the case of government consumption, fixed capital consumption traditionally accounts for a large share of total government consumption. However, the ratio of government purchases of goods and services has been increasing recently owing to the increase in expenditure on medical and elderly care. It is likely that this contributed to the recent structural changes in government consumption based on benchmark year 2015.

More salient features were observed in short-term fluctuations. In the case of GDP and government consumption, the 2015 benchmark revision brought about no conspicuous structural changes in short-term fluctuations for both subsample periods. For both GDP and government consumption, the data based on benchmark year 2015 showed peaks and troughs similar to those based on benchmark year 2011 over time. This implies that the 2015 benchmark revision did not change the short-term fluctuations in GDP or government consumption, even though it changed their levels or growth rates. Even for public investment, the benchmark revision did not change the short-term fluctuations for the first subsample, although it did change the level. However, it changed not only the level, but also the short-term fluctuations in public investment in the second subsample. Since the fourth quarter of 2011, short-term fluctuations in public investment have occasionally shown conspicuous discrepancies between the two benchmark years. For example, in 2015Q1, public investment in benchmark year 2015 declined sharply, whereas public investment in benchmark year 2011

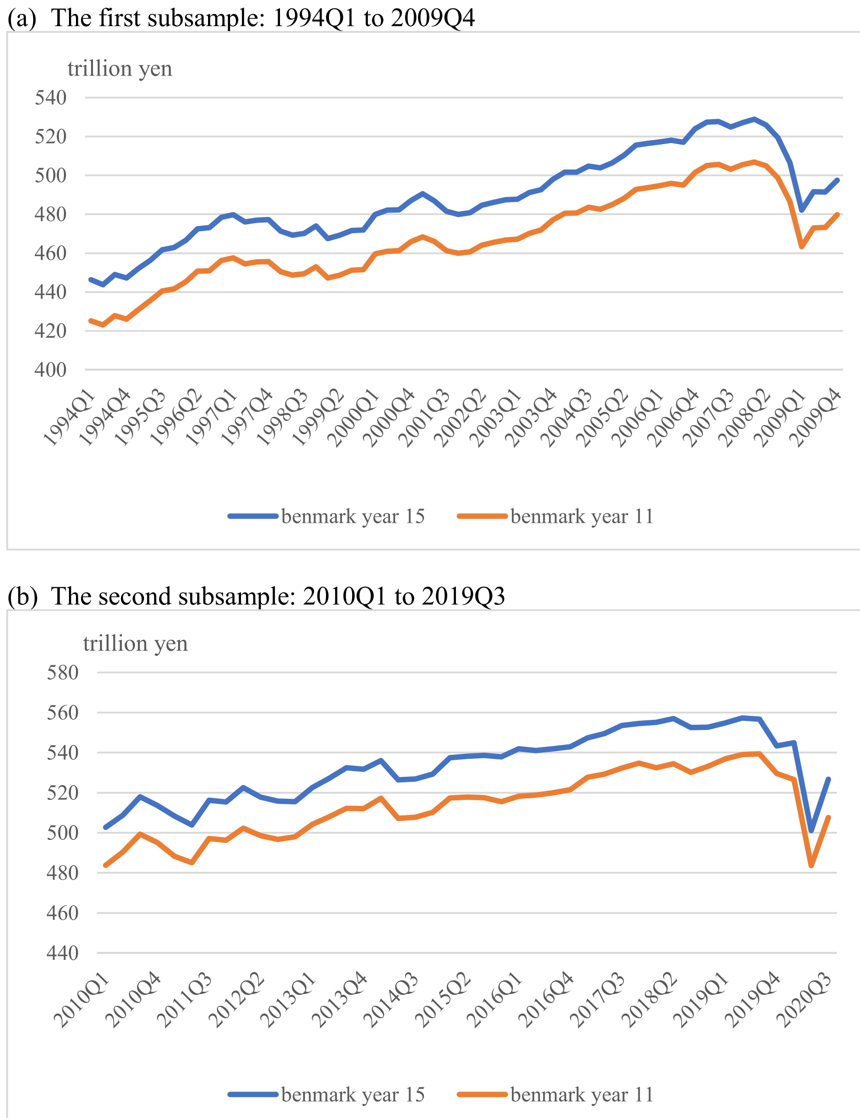

Fig. 6. Real GDP based on benchmark years 2011 and 2015.
Source: Quarterly Estimates of GDP (2008SNA), Cabinet Office.

declined only modestly. By contrast, from 2015Q3 to 2016Q1, public investment in benchmark year 2015 increased substantially, whereas that in benchmark year 2011 remained small. Instead, from 2016Q1 to 2016Q2, public investment in benchmark year 2015 declined substantially, whereas it increased modestly in benchmark year 2011.[4] This implies that the benchmark revision brought about a structural change in public investment because of the structural change in short-term fluctuations in 2015 and 2016.

The 2015 benchmark revision reflected the retroactive revisions of the Quick Estimate of Construction Investment by the Ministry of Land, Infrastructure, Transport, and Tourism since 2011. The retroactive revisions newly applied the progress rates in the “Construction Progress Rate Survey” and modified the correction rates in the statistics. As a result, they have changed the time-series properties of public construction data since 2011.[5] Because the Quick Estimate of Construction Investment is the key data of public investment in the SNA, it is likely that the 2015 benchmark revision caused a structural change in the short-term fluctuations of public investment since 2011.

The structural change in the short-term fluctuations in public investment was more serious in the quarterly data than in the annual data, especially in the fiscal year data. This is because the different timings in the quarterly data are smoothed out in the annual data and modified by the fiscal year data. Thus, it is likely that when using quarterly data after the structural break, the correlation coefficients between GDP and public investment became smaller for benchmark year 2015 than for benchmark year 2011. The structural change in short-term fluctuations, not a structural change in levels or growth rates, affected the effectiveness of fiscal policy and raised serious concerns in empirical studies in Japan.

[4] When looking at the changes in real public investment in real terms with calendar year data, the revisions show that in 2015 the year-on-year rate of decline has widened significantly. By contrast, the large expansion in the rate of the year-on-year increase in 2016 stands out.

[5] For example, the newly applied progress rates created a tendency that the time to complete construction would be longer than before in the data.

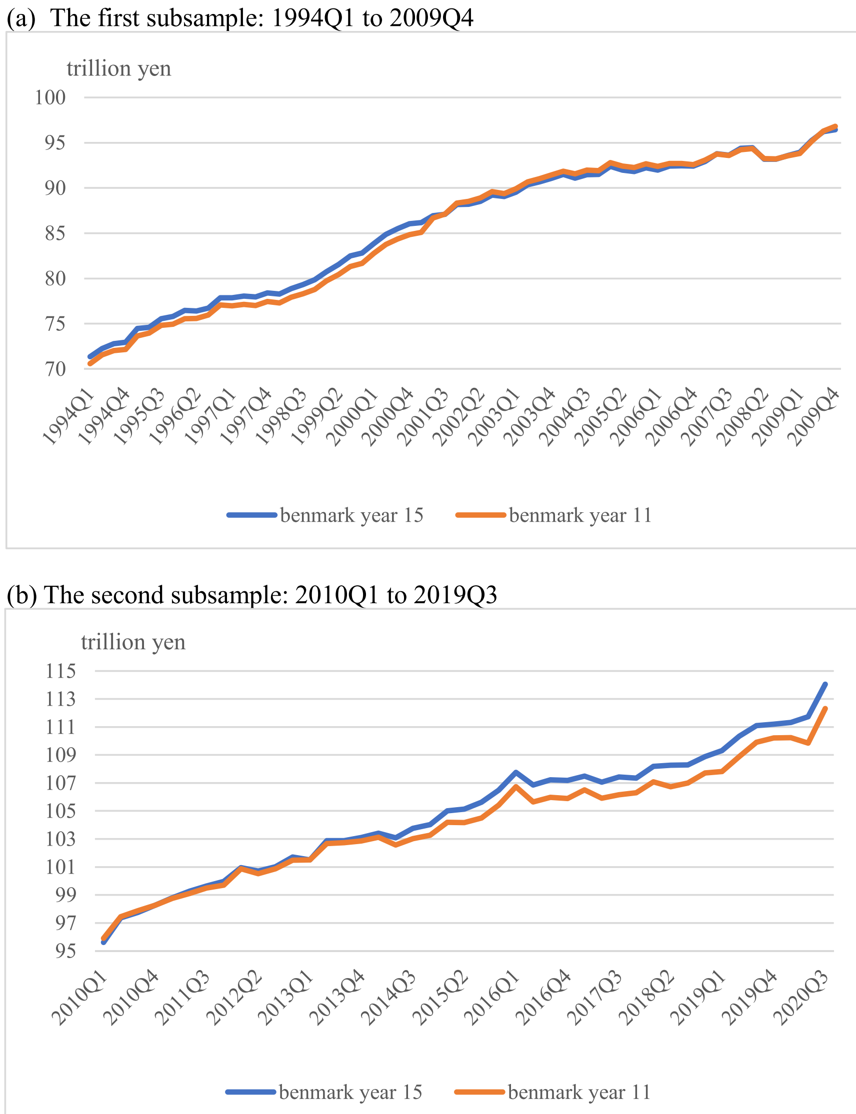

Fig. 7. Real government consumption based on benchmark years 2011 and 2015.
Source: Quarterly Estimates of GDP (2008SNA), Cabinet Office.

9.2. The discrepancy of each component of public investment

It is easy to confirm that different short-term fluctuations in public investment occurred in benchmark years 2015 and 2011, not only in real terms, but also in nominal terms. In the following analysis, we explore the components of the nominal public investment that exhibit different short-term fluctuations in the two benchmark years.

In Japan, the Annual Report on National Accounts publishes seasonally unadjusted nominal quarterly data on the Gross Capital Formation of the Private and Public Sectors. It divides the gross capital formation of the public sector into residential investment of public corporations, nonresidential investment of public corporations, gross capital formation of the general government, and changes in inventories of public corporations and the general government. Using these components, we examined which components showed different short-term fluctuations in public investment in benchmark years 2015 and 2011.[6]

Figs. 9, 10, and 11 show the quarterly time series of residential investment of public corporations, nonresidential investment of public corporations, and gross capital formation of the general government, respectively, based on benchmark years 2011 and 2015. Because the benchmark revision made short-term fluctuations in public investment quite different between 2015 and 2016, we depict time-series data from 2014Q1 to 2017Q4.

Because the data were seasonally unadjusted, we observed large seasonal fluctuations in all series. However, discrepancies were observed between the two benchmark years. The residential investment of public corporations based on benchmark year 2015 was smaller than that based on benchmark year 2011 in 2015Q2, 2016Q2, and 2017Q2 and larger than that based on benchmark year 2011 in 2014Q2, 2016Q1, 2016Q4, and 2017Q1. The non-residential investment of public corporations based on benchmark year 2015 was smaller than that based on benchmark year 2011 in 2014Q2, 2015Q1–2015Q3, 2016Q2–2016Q3, and 2017Q2–2017Q3, and larger than that based on benchmark year 2011 in 2014Q3 and 2016Q1. Gross capital formation of the general government based on benchmark year 2015 was larger than that based on benchmark year 2011 in 2016Q1 and 2017Q1.

These results suggest that nonresidential investments by public corporations show more frequent deviations between the two benchmark

[6] Because the share of inventories is negligible, the following discussions will skip examining short-term fluctuations of inventories of public corporations and general government.

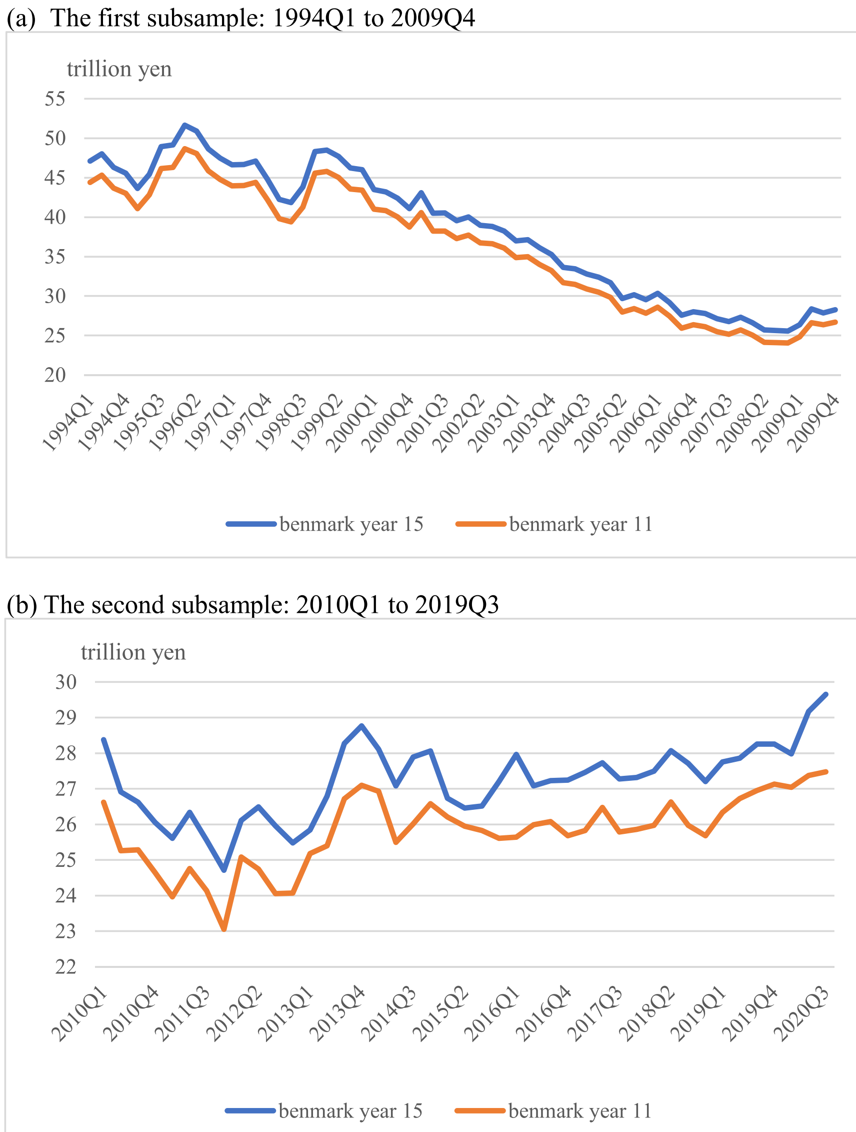

Fig. 8. Real public investment based on benchmark years 2011 and 2015.
Source: Quarterly Estimates of GDP (2008SNA), Cabinet Office.

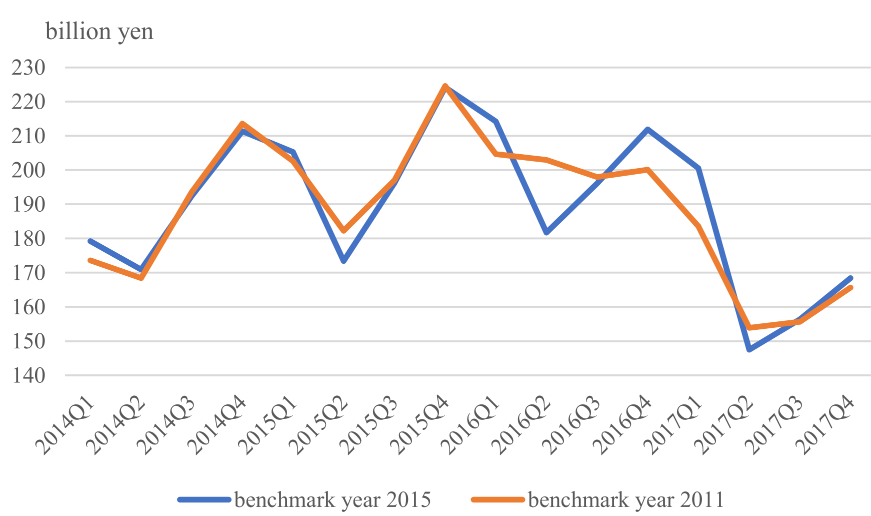

Fig. 9. Residential investment of public corporations.
Source: Annual Report on National Accounts, Cabinet Office.

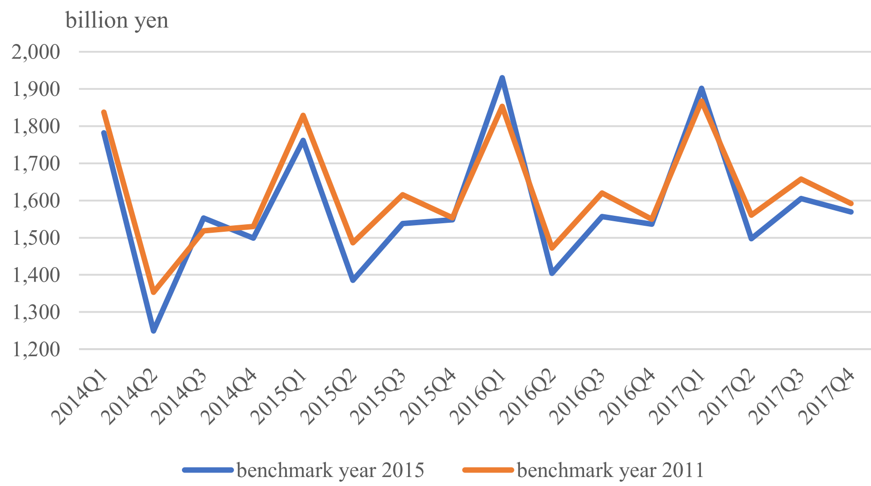

Fig. 10. Non-residential investment of public corporations.
Source: Annual Report on National Accounts, Cabinet Office.

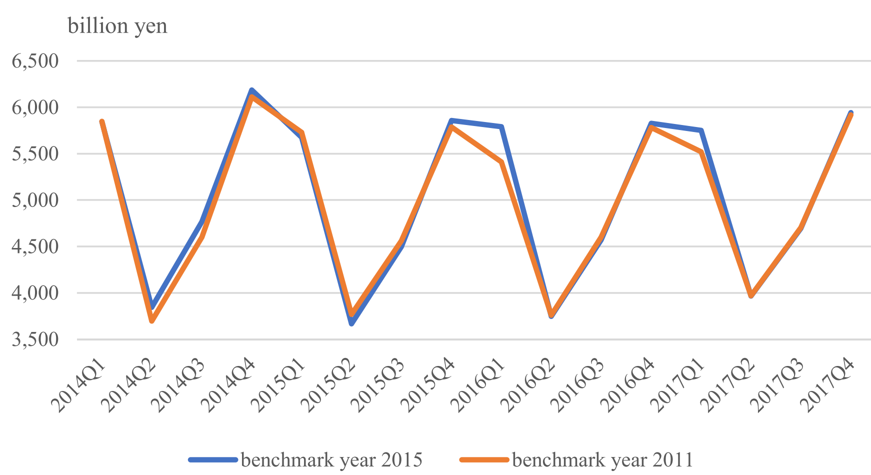

Fig. 11. Gross capital formation of general government.
Source: Annual Report on National Accounts, Cabinet Office.

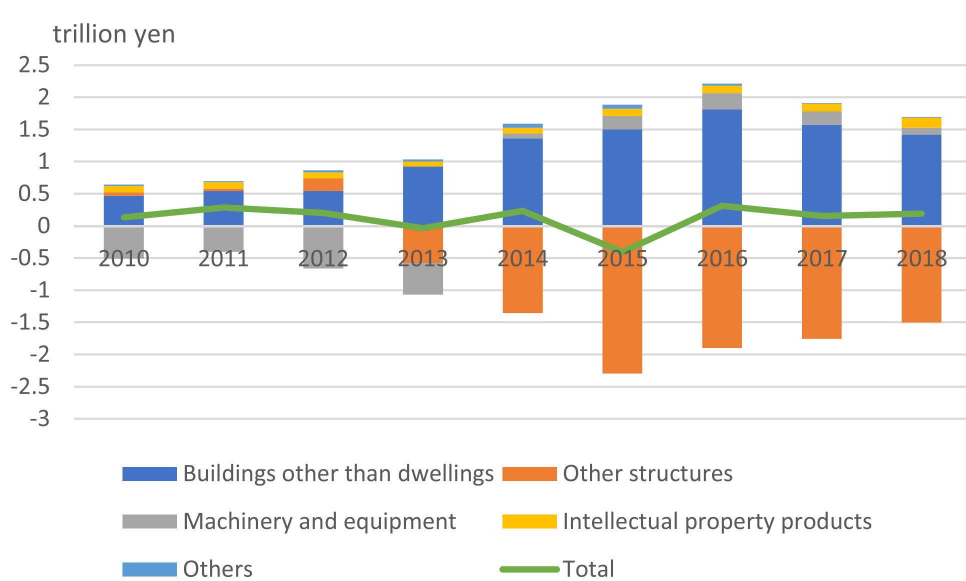

Fig. 12. Revisions of gross capital formation of public sectors.
Source: Gross Fixed Capital Formation of Assets classified by Institutional Sectors and Economic Activities (at current prices), Cabinet Office.

years. However, they also imply that each component of nominal public investment contributes to short-term discrepancies between the two benchmark years. For example, all components based on benchmark year 2015 were larger than those based on benchmark year 2011 in 2016Q1. Correspondingly, the real public investment based on benchmark year 2015 showed a sharp upward spike in 2016Q1, whereas that based on benchmark year 2011 did not. It is likely that the 2015 benchmark revision, which reflected the retroactive revisions of the Quick Estimate of Construction Investment since 2011, changed the short-term fluctuations of each public investment component and made the short-term fluctuations of aggregated public investment quite different. It is also likely that these changes in short-term fluctuations affected the effectiveness of the fiscal policy in our estimations.

The sources of the discrepancies were confirmed using the SNA annual estimate data. Using the Gross Fixed Capital Formation of Assets classified by Institutional Sectors and Economic Activities (at current prices), Fig. 12 depicts how the gross fixed capital formation of the public sector and its components in benchmark year 2015 deviated from those in 2011. This shows that the total gross fixed capital formation of the public sector was revised upward in 2014 and 2016 and downward in 2015 by the benchmark revision. More importantly, "buildings other than dwellings (i.e., building investment)" were always the main source of upward revision, whereas "other structures (i.e., civil engineering investment)" were always the main source of downward revision. Building investment was revised upward because renovation and renewal of buildings were included, whereas civil engineering investment was revised downward because the correction rates of the raw data were modified. In 2015, a large modification in correction rates caused a large downward revision in civil engineering investment, which contributed to the downward revision of public investment in benchmark year 2015. Correction rates can sometimes be large because of the low accuracy of the Prompt Report of Current Survey on Orders Received for Construction, which is the basic statistics for civil engineering investment. Because the construction periods and progress patterns of civil engineering and building investments differ, the quarterly pattern of change is expected to be more significantly affected by the benchmark revision than the annual pattern.

10. Concluding remarks

In this study, we examine the effectiveness of macro-fiscal policies using a GDP data series based on alternative benchmark years. The sample period used in the analysis includes the period of the "liquidity trap," in which interest rates fell to zero. It also included a period during which various structural problems, such as population aging, became conspicuous in the Japanese economy. The results suggest that the impact of fiscal policy on GDP might be quite different depending on GDP benchmark year we use in the analysis, especially for the sample period in which the 2015 benchmark revision reflected retroactive revisions of the public construction.

Studies have long debated the effectiveness of fiscal policies. However, the results of this study must be interpreted with caution, depending not only on GDP benchmark year but also on the choice of outliers and the estimation period. In the future, further verification is necessary to confirm the robustness of these findings on the effectiveness of fiscal policy in Japan.

Acknowledgments

An earlier version of this paper was presented at JWE conference "Government Statistics and Economic Performance" in Kyoto on March 26, 2022. I would like to thank the participants, especially Masahiro Higo, for their constructive comments. This research was supported by JSPS KAKENHI Grant Numbers 19H01466 and 20K20751.

References

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.
Auerbach, A.J., Gorodnichenko, Y., 2017. Fiscal multipliers in Japan. Res. Econ. 71, 411–421.
Bayoumi, T., 2001. The morning after: explaining the Slowdown in Japanese Growth in the 1990s. J. Int. Econ. 53, 241–259.
Bessho, S., 2021. Local fiscal multipliers and population aging in Japan. Jpn. World Econ. 60, 101099.
Blanchard, O., 2019. Public debt and low interest rates. Am. Econ. Rev. 109, 1197–1229.
Brückner, M., Tuladhar, A., 2014. Local government spending multipliers and financial distress: evidence from Japanese prefectures. Econ. J. 124, 1279–1316.
Caballero, R.J., Hoshi, T., Kashyap, A.K., 2008. Zombie lending and depressed restructuring in Japan. Am. Econ. Rev. 98, 1943–1977.
Fukuda, S., Nakamura, J., 2011. Why did 'zombie' firms recover in Japan? World Econ. 34, 1124–1137.
Fukuda, S., Yamada, J., 2011. Stock price targeting and fiscal deficit in Japan: why did the fiscal deficit increase during Japan's lost decades? J. Jpn. Int. Econ. 25, 447–464.
Futagami, K., Konishi, K., 2018. Dynamic analysis of budget policy rules in Japan. J. Jpn. Int. Econ. 50, 72–88.
Iwata, Y., 2011. The government spending multiplier and fiscal financing: insights from Japan. Int. Financ. 14, 231–264.
Kameda, K., 2014. What causes changes in the effects of fiscal policy? A case study of Japan. Jpn. World Econ. 31, 14–31.
Kameda, T., Namba, R., Tsuruga, T., 2021. Decomposing local fiscal multipliers: evidence from Japan. Jpn. World Econ. 57, 101053.
Kato, A., Miyamoto, W., Nguyen, T.L., Sergeyev, D., 2018. The effects of tax changes at the zero lower bound: evidence from Japan. AEA Pap. Proc. 108, 513–518.
Kuttner, K.N., Posen, A.S., 2001. The Great recession: lessons for macroeconomic policy from Japan. Brook. Pap. Econ. Act. 2001, 93–185.
Kuttner, K.N., Posen, A.S., 2002. Fiscal policy effectiveness in Japan. J. Jpn. Int. Econ. 16, 536–558.
Miyamoto, W., Nguyen, T.L., Sergeyev, D., 2018. Government spending multipliers under the zero lower bound: evidence from Japan. Am. Econ. J. Macroecon. 10, 247–277.
Morita, H., 2015. Japanese fiscal policy under the zero lower bound of nominal interest rates: time-varying parameters vector autoregression. IER Discussion Paper A No. 627. Hitotsubashi University.
Otsu, K., Shibayama, K., 2022. Population aging, government policy and the postwar Japanese economy. J. Jpn. Int. Econ. 64, 101191.
Saxegaard, E.C.A., Davis, S.J., Ito, A., Miake, N., 2022. Policy uncertainty in Japan. J. Jpn. Int. Econ. 64, 101192.
Umeda, M., Kawamoto, T., Hori, M., 2018. Twin survey on the Japanese economy and policy effects: outline of the survey and report of summary statistics. Econ. Anal. 197, 144–185.
Werner, R., 2004. Why has fiscal policy disappointed in Japan? In: Proceedings of the Money Macro and Finance (MMF) Research Group Conference. Money Macro and Finance Research Group, p. 9.
Yoshino, N., Miyamoto, H., 2017. Declined effectiveness of fiscal and monetary policies faced with aging population in Japan. Jpn. World Econ. 42, 32–44.