RUHR

ECONOMIC PAPERS

Roland Döhrn

Transportation Data as a Tool for Nowcasting Economic Activity

The German Road Pricing System as an Example

Imprint

Roland Döhrn

Transportation Data as a Tool for Nowcasting Economic Activity

The German Road Pricing System as an Example

Bibliografische Informationen
der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen National-
bibliografie; detaillierte bibliografische Daten sind im Internet über:
http://dnb.d-nb.de abrufbar.

Roland Döhrn[1]

Transportation Data as a Tool for Nowcasting Economic Activity – The German Road Pricing System as an Example

Abstract

There is a broad agreement that transportation activity is closely linked to the business cycle. Nevertheless, data from the transportation sector have not been part of the tool kit of business cycle analysts due to long publications lags. With the disseminations of electronic road pricing systems, up to date figures on transportation activity are available for an increasing number of countries. This paper analyses the performance of the German toll statistics for nowcasting industry production. It confirms that between January 2007, when the toll data were published first, and July 2012 the seasonally adjusted toll data show a closer correlation with industry production than business surveys like the ifo business climate or the PMI. Compared to this the forecasting power out of sample is disappointing. Though showing somewhat smaller forecast errors than the alternative models tested the advantage of the toll based models is not statistically significant as a rule. Given the small publication lead against industry production and the publication lag against business sentiment indicators one should not be over-enthusiastic on the opportunities of the toll data as a nowcasting tool, though they surely mean an addition to the business cycle analysts' tool box.

JEL Classification: E32, E37

Keywords: Transportation data; nowcasting; forecasting performance

January 2013

1 Introduction

As Lahiri/Yao/Young (2004a,b) have pointed out, transportation activities react very sensitively to fluctuations in the business cycle. Nevertheless, indicators derived from the transportation sector have not been a part of the tool kit of most business cycle analysts for many years. An important reason for neglecting this kind of data was probably their long publication lag. However, in recent years the availability of such data has improved markedly. One reason is that more and more countries use electronic systems to collect road tolls, which make access to data on transportation activities much easier and reduce the publication lag considerably. Austria, e.g., introduced an electronic road pricing scheme for trucks, busses and motor homes already in 2004, the Czech Republic in 2007 and Slovakia in 2010. In Austria data are available only a few days after the end of each month which are used by the Austrian National Bank (OeNB) to calculate a leading indicator of Austrian exports (Fenz, Schneider 2009).[2]

Germany introduced its toll system in 2005. Starting in 2008, the Federal Office for Goods Transport *(Bundesamt für Güterverkehr; BAG)* publishes several monthly indicators derived from this system[3], which are available as a rule 15 working days after the end of a month, i.e. earlier than the corresponding figures on industry production. This source has attracted the attention of researchers in the meantime. Askitas and Zimmermann (2012) proposed a toll index which, in their own assessment, provides "a reliable early nowcast of the state of the economy in general" (Ibid 2011: 2). However, their work suffers from two shortcomings: Firstly they concentrate on figures which have not been seasonally adjusted, although the transportation activity displays a strong seasonal pattern, which should be taken into account. Secondly,

---

[2] The Austrian data are only made available to the OeNB and not accessible for the public.

[3] As the publication also contains a comparison to the situation one year earlier, the data in fact go back to 2007.

they do not evaluate their index against other widely used indicators such as business surveys.

By contrast, Döhrn (2011) focused on seasonally adjusted data. On a positive note, he confirmed the results of Askitas and Zimmermann concerning the strong correlation of the toll data with economic activity. But he also found reasons for skepticism about the opportunities the data offer for business cycle analyses. The toll data are published later than the survey data for the corresponding month. The question is, whether this disadvantage can be compensated by a closer correlation with industry activity. In the period analyzed, at least, forecasts of industry production based on toll data did not outperform forecasts using the ifo business climate index. However, these first analyses suffered from the brevity of the time series available, since the monthly data only go back until 2007. This problem was aggravated by the fact that the data cover a period in which the strong decline of economic activity during the deep recession 2008/09 was dominant. As each indicator of economic activity was depressed in these years, the correlation between all of them appeared to be high.

Some of the problems of earlier studies can be cured when the more data are available. With time series getting longer, the 2008/09 recession will lose its influence, and it will become clearer, what the toll data can tell beyond identifying the recession. Furthermore, the seasonal adjustment procedure can be expected to generate more stable results. However, in August 2012 there was a break in the data, since the liability to pay a toll was extended to other types of roads. Therefore, only data until July 2012 can be used in this study. Thus the possibilities to get clearer results by having more data are limited. To improve the quality of the data anyway, a new method for backcasting data was employed which will lead to a more stable estimate of the seasonal pattern and therefore generate better seasonal adjusted figures.

The paper is organized as follows: The next section the toll data will be presented and the problems associated with the estimation

of seasonally adjusted figures will be discussed. In section 3, a regression analysis of the toll data in relation to economic activity will be conducted and its results will be compared to regressions using survey based indicators. Then, out of sample forecasts based on these estimates will be evaluated. Finally, some conclusions will be drawn on the future role of toll data in business cycle analyses.

2   A seasonally adjusted series of toll data

The road toll on highways and on other selected four lane roads in Germany has to be paid for transportation by trucks with a permissible total weight of 12 tons and above. The toll system offers a manual charging at toll station terminals or via the internet, and an automatic charging. The latter is based on a combination of mobile radio technology (GMS standard) and the global positioning system (GPS). Trucks using electronic charging must have an On Board Unit (OBU) that allows tracing the route a truck takes via satellite (for details BAG 2011). From the material gathered by these techniques BAG calculates four indicators that are linked to economic activities: (i) the number of kilometers driven on roads that are subject to the toll. (ii) the number of toll rides (iii) the number of toll vehicles crossing the German border inbound, and (iv) the number of toll vehicles crossing it outbound[4]. The two last ones will be neglected subsequently as they show a poor correlation with economic activity in Germany (Askitas, Zimmermann 2011), which might reflect a high share of transit traffic.

All data are published timely, and they are subject to only very small revisions. Both characteristics form good preconditions that the data may serve as short term indicators. Nevertheless, the adequacy of the indicators has to be scrutinized, as the relation between toll rides and economic activity may be influenced by

---

[4]In addition, data on environmental issues are provided. Furthermore, all data are disaggregated according to the national origin of the vehicles.

reactions of those who are liable to the toll. Thus, carriers might try to avoid the toll by using smaller trucks or roads which are not subject to the toll. This might be more relevant in business cycle downturns, thus distorting any forecast built on toll data. Furthermore, the laws defining the tax base has been altered from time to time. In August 2012, e.g., the liability to pay the toll was extended to other types of roads. Another conceivable extension could be to include vehicles with a permissible total weight below 12 tons.

For good reasons, business cycle analysts are mostly interested in seasonally adjusted figures. This kind of data, however, cannot be observed directly, but must be estimated. Therefore the results may vary considerably with respect to the method employed and the data used. However, the alternatives to seasonal adjustment are not very satisfying. Month over month variations of unadjusted data are strongly influenced by the seasonal pattern, and it is difficult to discriminate between seasonality and the information relevant for assessing the current situation. Year over year variations, which form the basis of the estimates of Askitas/Zimmermann (2011, 2012), necessarily generate a lagging indicator, because they represent the accumulated change over the past twelve months, and not the most recent developments.

However, seasonally adjusted figures are particularly difficult to determine when time series are short. If the data at hand cover less than three years, no seasonal adjustment should be conducted at all, and for some algorithms it is even impossible from a technical point of view. For time series between three and seven years, EUROSTAT (2009: 29) recommends to check the robustness of the seasonal adjustment parameters very carefully. With time series being short, each new data point adds a considerable piece of information to the adjustment process, whatever technique is used. In the case considered here, this problem is aggravated by the fact that the 2008/09 recession reached its trough during January and February 2009, i.e. in two months in which according to the normal seasonal pattern transportation activities are low

anyway. This makes it even more difficult to discriminate between seasonal and cyclical effects.

Furthermore, Matas-Mir/Rondonotti (2003) have shown that for the Census-X12-method, which is employed here, outliers at the beginning of a sample have a particularly strong impact on seasonal adjustment results, because they influence the ARIMA backcast of the time series to be adjusted, which is necessary to get adjusted figures for the entire sample. This problem can be overcome at least partially by replacing the automatic ARIMA backcast of CENSUS method by an explicit backcast using a priori information. Here the fact can be utilized that the toll data and industry production display a similar seasonal pattern. Therefore, non-adjusted levels of monthly industry production in the years 2005 and 2006 are used to estimate hypothetical values of the toll data. To be consistent, industry production is related to the number of working days in the same way as the toll data. Furthermore, the results are controlled for a variable representing the weather conditions. Since the backcasted figures are only used for the seasonal adjustment and do not enter the analyses of the interrelation of toll data and production activity, this procedure does not spoil the results of our subsequent analyses.

Including this explicit backcast, the seasonal adjustment is conducted in two steps. Firstly, in a pre-treatment the data are corrected for a working-day effect. The calculations are based on information provided by BAG together with the toll data, according to which transportation activity reaches only 30% of the activity of a normal working day on Saturdays and only about 11% on Sundays. This information is used to standardize the number of working days per month and to calculate the transportation activity per standard working day. In a second step, these series are seasonally adjusted using the default settings of CENSUS X12-Arima implemented in Eviews®.

To check as recommended the robustness of the results, the seasonal adjustment is carried out stepwise. Initially, the calculation is made for the period Jan 2005 to Jan 2010. Then the

sample is repeatedly extended by one month and the adjusted figures are calculated again. Thus, all in all 26 series of seasonally adjusted data are generated. The robustness of the results is measured by the mean absolute change of the adjusted figure from step t to step t+1. In Chart 1, this robustness measure is presented for the kilometers travelled. It exemplifies that indeed the inclusion of each additional data point leads in a somewhat different picture of the seasonally adjusted development. However, the series generated by using the explicit backcast seem to be more robust than those generated by help of the ARIMA backcast. This particularly holds for the months at the beginning of the sample, for which the backcast problem weighs most. Of course, the end-of-sample problem in determining the underlying trend cannot be overcome as easily, so that robustness measures for the more recent months do not differ very much between the two adjustment techniques.

The unadjusted and the adjusted time series are provided in chart 2 for the number of trips and the kilometers travelled. Some large month to month variations in the adjusted data suggest that there is still a strong irregular component in the data which can reduce the forecasting power of the indicator. Therefore, the trend-cycle component of the seasonally adjusted data is estimated additionally. It shows a smoother profile, but the information content of the most recent data may be reduced by this transformation.

3 Estimation results

The seasonally and working day adjusted series estimated from monthly data for the period January 2007 to July 2012 will be used now to make some one step forecasts of industry production. For that purpose quite a number of regression models have been tested. The left hand side variable always is the month over month variation of seasonal and working day adjusted industry production. On the right hand side parsimonious specifications have been chosen to save degrees of freedom for estimation (table 1). In (1), the month over month variation of the kilometers

travelled is used as the only explanatory variable. Model (2) is motivated by the finding of Askitas/Zimmermann (2012) according to which the explanatory power can be improved by using the number of trips as an additional variable. In (3) model (1) is supplemented by an ARMA(1,1) term.

The next three models (4) to (6) are similar to specifications (1) to (3), but they use the trend-cycle components of the explanatory variables instead to eliminate the impact of the irregular component. The next three models in table 1 are thought to serve as a yardstick for the toll indicator models. Model (7) is an ARMA(1,1)-model. Model (8) uses the ifo business climate indicator as an alternative to the toll data, and model (9) the Purchasing Managers' Index (PMI) for the manufacturing sector. These two survey based indicators are published about four weeks earlier than the toll data. Specifications (10) to (13) finally use the toll data additionally to the business climate and PMI respectively to test whether the toll data contain information which is not already in the surveys.

Table 1 shows the in-sample performance of these models. It makes evident that there is a statistical significant correlation between the toll kilometers and industry production in all specifications. Compared to Döhrn (2011), the correlation is lower now for the larger sample which underpins that the toll data have been a powerful indicator particularly during the Great Recession. But also for the period analyzed here, the trend-cycle component generates a better fit than the original seasonally adjusted data. Complementing the kilometer variable by the number of trips or an ARMA(1,1) term does not improve the results. Compared to the reference models, the toll data outperform the ARMA (1,1) model and the ifo business climate. The PMI performs better than the seasonally adjusted kilometers, but it does not better than the trend-cycle component of the toll kilometers. The picture appearing when the toll data are combined with the business survey figures is mixed: Equations (10) and (11) which include the seasonally adjusted kilometers travelled and the survey based

indicators show a better fit than the models using only one of the variables as the only regressor. However, in combinations with the trend-cycle component (equations 12 and 13), business climate as well as PMI become insignificant. All in all, the in-sample performance suggests that the trend-cycle component of k could be a powerful predictor of economic activity in industry.

4   Forecast performance

However, a good fit in sample does not necessarily guarantee a superior forecast performance out of sample. As a starting point, the toll data were seasonally and working-day adjusted for the period January 2007 to January 2010 (using the explicit backcast to January 2005), and the models (1) to (13) were estimated for this sample. Then the change of industry production in January 2010 was forecasted. Thereafter, the sample was expanded by one month and the next forecast was made. This procedure was repeated until the sample ended in July 2012 forecasting industry production for the same month. Thus, 31 forecasts of the month over month change of industry production were generated.

Table 2 presents a forecast evaluation for those models in table 1 in which the coefficients of the explanatory variables have been significant. As a measure of forecast accuracy the mean squared errors (MSE) of the forecasts are presented. Furthermore, the small sample variant of the Diebold-Mariano test suggested by Harvey, Leybourne, Newbold (1997) is shown for a comparison of the forecast error with the ARMA (1,1) model (MDM). The forecast errors were calculated in two ways. Firstly, the first release data of month over month variation of industry production were taken as a yardstick. Secondly, the forecasts were related to the latest available data. By doing so, it is taken into account that industry production often is revised substantially while the toll data are subject to very small revisions only. Thus the toll data may be a better predictor of the final industry production data rather than of the preliminary results.

Given the superior performance of the toll-based models in-sample, the out of sample properties are somewhat disappointing. Compared to the first release data, their MSE is only slightly smaller than the forecast error of the three reference models. However, as the MDM statistic shows, none of the model (1) performs significantly better than the ARMA (1,1) model. But the same is true for model (8) (using the ifo business climate instead of the toll data) and the mixed models (10) and (11). Comparing the forecast with the latest available data on industry production reduces the forecast error, which indicates that the toll data indeed may help to forecast revisions of industry production. However, the same is true for all other models considered, so that the toll data do not perform significantly better than other indicators or even an ARMA (1,1) model.

However, the forecast error covers only one dimension of the information provided by the toll data and its competitors. Given the high volatility of the seasonally adjusted rates, forecasters may also be interested in knowing the dynamics of industry production, here understood in the sense whether growth is going to speed up or to cool down. To test for this feature of forecast accuracy a measure of co-movement is applied which was suggested by Harding/Pagan (2002). For calculating it, the time series of growth rates are in a first step transformed into binary state variables which are defined as

$$S_{r,t} = \begin{cases} 1 & if \quad dlog(ip_t) - dlog(ip_{t-1}) > 0 \\ 0 & else \end{cases}$$

$$S_{f.t}^{x} \begin{cases} 1 & if \quad dlog\left((ip_{f,t}|x)\right) - dlog(ip_{t-1}) > 0 \\ 0 & else \end{cases}.$$

S_r describes changes in the dynamics of the realized growth rates. It is one if the rate of change is accelerating and zero when it is decelerating. S_f^x is the corresponding measure for the forecasted changes using indicator x as a predictor. In the variant used here the forecasted growth is related to the last observation.

Based on these definitions a concordance indicator I_x is determined as

$$I_x = \frac{1}{T}\left(\sum_{t=1}^{t} S_{r,t}S_{f.t}^{x} + \sum_{t=1}^{t}(1 - S_{r,t})(1 - S_{f.t}^{x})\right).$$

For testing the significance of I_x Harding/Pagan (2006) propose a simple t-test which is based on a regression of $S_{r,t}$ on $S_{f,t}^{x}$, after dividing the two variables by their standard deviation. Following Lahiri/Yao/Young (2003) a Newey-West heteroscedasticity and autocorrelation consistent estimator is used for running this test.

Table 3 shows that all models employed provide information on production dynamics. The measure of concordance is between 0.77 and 0.87 and it is significant in all cases. The highest value is observed for the model using seasonally adjusted number of kilometers, the lowest value for PMI. However, the ARMA (1,1) model shows almost the same performance as the competing models.

5   Conclusions

This paper assesses the power of the toll data as a tool for nowcasting German economic activity in the case seasonally adjusted figures are used. In sample, the number of kilometers travelled by vehicles subject to the road toll is a better forecaster of industry production than an ARMA (1,1) process, and it also outperforms survey based business climate indicators. Extending the model by including the number of trips or autoregressive terms as additional variables does not increase the fit further. The same is true when combining toll data with the surveys. Out of sample, the toll based models generate more accurate forecasts than a simple ARMA (1,1), but their forecasts are not significantly better than those using survey based indicators. The seasonally adjusted number of kilometers travelled also seems to be a good forecaster of production dynamics. But again, the advantage against other models is small.

Against this background, toll data surely provide additional insights to business cycle analysts. However, one should not be too enthusiastic about the opportunity this newly available source offers. Firstly, the publication lead of toll data vis a vis industry production is small, and the data are published later than well-established indicators such as business surveys. This disadvantage is not compensated by a superior forecast performance. Secondly, toll data will be influenced by changes in the regulatory environment. Most recently, the toll road network has been extended by about 1100 km.[5]. Thus, the future results of the toll statistics will not be fully comparable to the data we have now, which will reduce their power as a forecasting device.

References:

Askitas, N and K.F. Zimmermann (2011), The Toll Index: Innovation-based Economic Telemetry. IZA Policy Paper No. 31. Bonn.

Askitas, N. and K. F. Zimmermann (2012), Nowcasting Business Cycles Using Toll Data. Journal of Forecasting. DOI:10.1002/for.1262.

BAG (2011), Toll Statistics. Methodological Explanations. http://www.bag.bund.de/SharedDocs/Downloads/DE/Statistik/Lkw-Maut/Meth_Erlaeuterung_Maut_englisch.pdf?__blob=publicationFile. Downloaded on 6 February 2012.

Döhrn, R. (2011), Die Mautstatistik: Keine „Wunderwaffe" für die Konjunkturanalyse. Wirtschaftsdienst 91: 863-868.

EUROSTAT (2009), ESS Guidelines on Seasonal Adjustment. Luxembourg.

Fenz, G. and M. Schneider (2009), A Leading Indicator of Austrian Exports Based on Truck Mileage. Monetary Policy & Economy Q1/09: 44-52.

Harding, D. and A. Pagan (2002), Dissecting the cycle: a methodological investigation. Journal of Monetary Economics 49: 365-381.

---

[5] For comparison reasons BAG announced to provide for one year additionally data on toll kilometers for the toll road network without the extension having come into effect on August 1st 2012..

Harding, D. and A. Pagan (2006), Synchronization of cycles. Journal of Econometrics 132: 59-79.

Harvey, D., S. Leybourne and P. Newbold (1997), Testing the equality of prediction mean squared errotrs. International Journal of Forecasting 13: 281-291.

Lahiri, K., W. Yao and P. Young (2004a), Cycles in the Transportation Sector and the Aggregate Economy. University of Albany Discussion Papers 03-14.

Lahiri, K., W. Yao and P. Young (2004b), Transportation and the Economy. Linkages at Business-Cycle Frequencies. Tranportation Research Record 1864: 103-111.

Matas, Mir, A. and V. Rondonotti (2003). The performance of X-12 in the seasonal adjustment of short time series. In: M. Manna and R. Peronaci, Seasonal adjustment. Frankfurt am Main, ECB: 149-159.

Table 1: Regressions on month over month variations of seasonal- and working day adjusted industry production; Feb 2007 to March 2012

	Constant	Kilo-meters	Trips	Business climate	PMI	AR(1)	MA(1)	R² adj.
(1)	0.00100	0.482**						0.147
(2)	0.00111	0.651*	-0.160					0.145
(3)	0.00055	0.355**				0.755**	-0.606	0.152
(4)	0.00170	1.722***						0.374
(5)	0.00231	2.798*	-1.408					0.374
(6)	0.00169	1.724***				-0.360	-0.144	0.384
(7)	-0.00050					0.774***	-0.545	0.085
(8)	-0.00127			0.00034***				0.121
(9)	-0.07062				0.00138***			0.254
(10)	-0.00058	0.404**		0.00011*				0.215
(11)	-0.05864	0.312*			0.00115***			0.304
(12)	0.00137	1.644***		0.00004				0.366
(13)	-0.01775	1.441***			0.0004			0.373

Notes: Abbreviations see text. In (4)-(6) and (12)-(13) the trend-cycle components of k and t are used. *p<0.05; **p<0.01; ***p<0.001

Table 2: Out of sample performance of the forecasting models

Model	MSE	MDM[1]	MSE	MDM[1]
	First release industry production		Latest available industry production	
(1)	2.71	2,4*	2.29	0.4
(4)	3.29	1.7	2.04	1.2
(7)	4.08	-	2.60	-
(8)	3.47	2.1*	2.19	2.1*
(9)	3.08	2.2*	1.73	2.4*
(10)	2.71	2.1*	2.07	0.8
(11)	2.61	2.1*	1.79	1.2

Note: For abbreviations see text, for the specification of the models table 1.
**:p<0,05. – [1]Relative to the ARMA (1,1)-model.*

Table 3: Concordance of the predicted and the observed dynamics in industry production

Model	I	t-statistics
(1)	0.86	8.6***
(4)	0.79	5.2***
(7)	0.79	4.2***
(8)	0.75	4.0***
(9)	0.75	4.2***
(10)	0.82	7.9***
(11)	0.82	7.9***

*Note: For abbreviations see text, for the specification of the models table 1.*
***:p<0,001.

Chart 1: Robustness of the seasonally adjusted data: Explicit backcast versus ARIMA backcast
Mean absolute deviation caused by extending the sample by one month
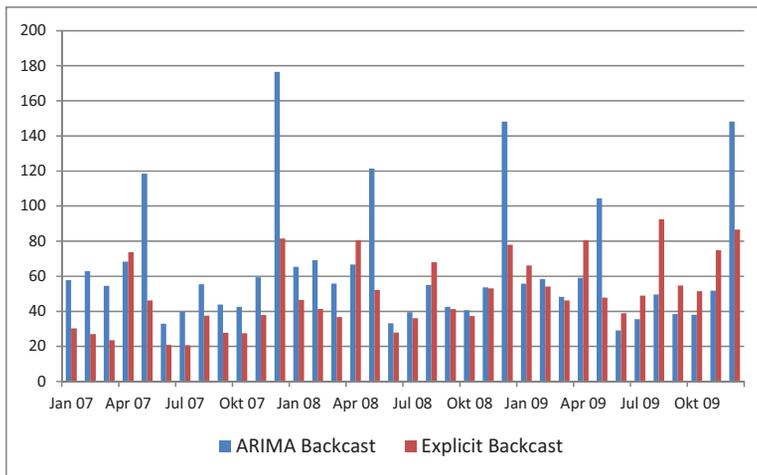
Author's calculations

Chart 2: Kilometers on toll roads and the number of toll rides
per working day

Kilometers (1000)

Toll rides
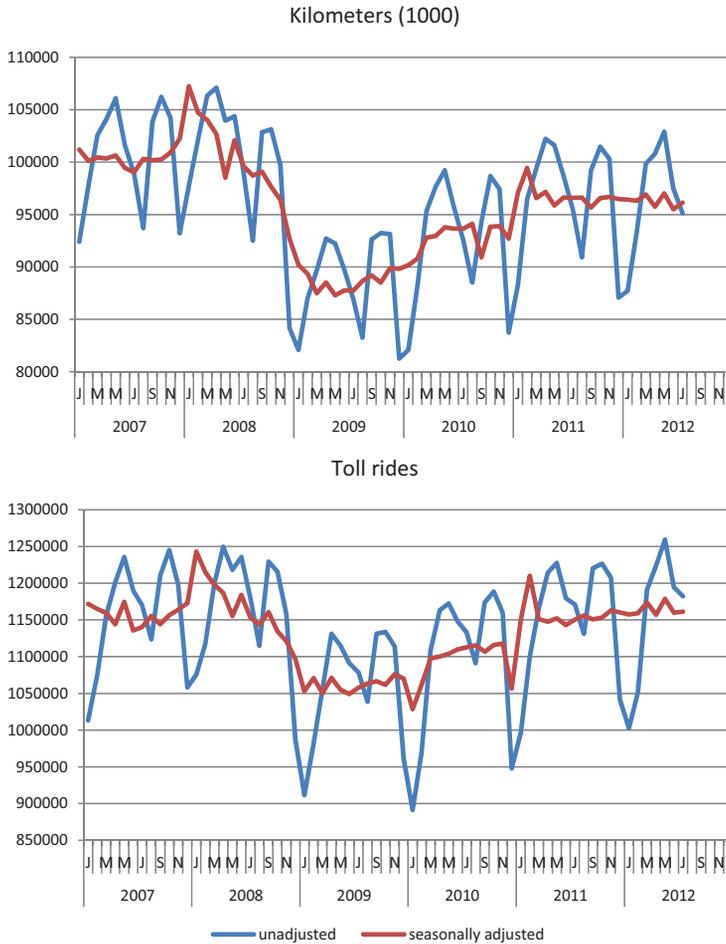
Author's calculation. Seasonal adjustment: CENSUS X12 with explicit backcast.