Policy impacts estimates are sensitive to data selection in empirical analysis: evidence from the United States – Canada softwood lumber trade dispute
Daowei Zhang and Rajan Parajuli

Abstract: In this paper, we use the U.S. softwood lumber import demand model as a case study to show that the effects of past trade policies are sensitive to the data sample used in empirical analyses. We conclude that, to be consistent with the purpose of analysis of policy and to ensure all else being equal, policy impacts can only be judged by using data up to the time when the policy is terminated.

Key words: policy analysis, time-series data, time element, U.S.–Canada, softwood lumber trade dispute.

Résumé : Dans cet article, nous utilisons la demande de bois d’œuvre résineux importé par les États-Unis comme étude de cas pour montrer que les effets des politiques commerciales passées dépendent de l’échantillon de données utilisé dans les analyses empiriques. Nous arrivons à la conclusion que pour être cohérent avec l’objectif de l’analyse de politique et s’assurer que toutes choses sont égales par ailleurs, les impacts d’une politique ne peuvent être évalués qu’en utilisant des données qui s’échelonnent dans le temps jusqu’au moment où cette politique prend fin.

Mots-clés : analyse de politiques, série temporelle de données, élément temporel, différend canado-américain sur le commerce du bois d’œuvre résineux.

Introduction
Economists and policy analysts attempting to assess the impacts of public policy often use time-series regression methods. Sometimes, the time-series data they use are very long and cover several periods of distinguishable policy regimes. If they use dummy variables to measure the effects of various policy regimes, the data sample period they choose may influence the estimation results and hence the impacts of policy regimes. Further, the dummy variables representing the earlier policy regimes could even have the wrong sign and be significant, contrary to economy theory (Nagubadi and Zhang 2013; Parajuli and Zhang 2016).

The purpose of this paper is to demonstrate that estimates of policy impacts vary with the data sample used in empirical analyses. We use the softwood lumber trade dispute between the United States (U.S.) and Canada, which has experienced four separate policy regimes, as a case study, and show that the appropriate estimates for the earlier policy regimes are to use data up to the end of the regimes, not data afterwards. This paper is inspired by findings in Nagubadi and Zhang (2013) and Parajuli and Zhang (2016). To estimate an empirical model, we use the same theoretical framework and data employed by Nagubadi and Zhang (2013) and Parajuli and Zhang (2016). The next section presents briefly the trade dispute and the four different policy regimes, followed by empirical models and results. The final section concludes with discussion.

The four separate policy regimes in the U.S.–Canada softwood lumber trade
Zhang (2007) presented a detailed historical chronology, political economy, and explanation of the long U.S.–Canada softwood lumber trade dispute. Canada is the largest exporter of softwood lumber to the U.S. Free trade on softwood lumber had prevailed for several decades prior to 30 December 1986, when the two countries signed the Memorandum of Understanding (MOU) that started to limit Canadian exports to the U.S. via a 15% export tax or stumping adjustment on the part of Canadian provinces. The MOU lasted for nearly 5 years, until Canada withdrew in October 1991. This was the first policy regime in the largest and longest-standing dispute between the two countries. The MOU and other trade agreements between the two countries excluded the maritime provinces of Canada. Thus, similar to others who conducted empirical studies on this area (e.g., Wear and Lee 1993; Zhang 2001, 2006), we only include the restricted Canadian provinces. There were varying details in the number of provinces covered and among the provinces covered in all of the policy regimes. However, as we are focusing on the aggregate impacts and these variations are minor, they should not impact our results much.

The two countries then signed the Softwood Lumber Agreement in principle in February 1996, which became a formal trade agreement (henceforth referred to as SLA 1996) for 5 years. SLA 1996 expired on 31 March 2001 and was the second policy regime. Between August 2001 and October 2006, the U.S. government imposed countervailing and antidumping duties (CVDAD) on Canadian softwood lumber imports with a combined rate of tariffs ranging from 10.2% to 27.2%. This policy regime (CVDAD) ended on October 2006 when the two countries signed another Softwood Lumber Agreement (SLA 2006) that lasted until October 2015. Figure 1 depicts a graphical overview of Canadian lumber exports to the U.S. under the four trade policy regimes.
As the softwood lumber trade dispute is the largest trade dispute between the countries with some $4–7 billions of goods annually being involved, economists have studied the welfare impacts of various policy regimes or trade restriction measures. Wear and Lee (1993) and Myineni et al. (1994) found that MOU was effective in reducing Canadian lumber imports to the U.S., which benefited U.S. producers and harmed U.S. consumers. Likewise, Lindsey et al. (2000) and Zhang (2001, 2006) reached the same conclusion with respect to SLA 1996. Devadoss (2006), Mogus et al. (2006), and Song et al. (2011) found that CVDAD had significant negative impacts on Canadian lumber imports. Recently, Baek (2012), Nagubadi and Zhang (2013), and Parajuli and Zhang (2016) looked into the effects of SLA 2006 in its 7–9 years of operation on the U.S. lumber imports from Canada and found it to be effective.

All these studies have used time-series regressions to estimate empirical models of Canadian lumber exports to the U.S. In the regression used by Nagubadi and Zhang (2013) and Parajuli and Zhang (2016), where monthly data from 1980 to 2012 (or 2015) were used, the dummy variables used representing MOU and SLA 1996 were no longer significant or even had wrong signs. Baek and Yin (2006) even concluded that SLA 1996 was not effective. We think that this conclusion is not warranted for several reasons. First, it is contrary to the empirical evidence that, in each of the 5 years under SLA 1996, Canadian exporters surpassed the duty-free quota by paying an export tax of $100 per thousand board feet, which was more than 20% of the prevailing lumber prices in these years (Zhang 2007, pp. 144–145). Second, and as noted earlier, other empirical studies (e.g., Zhang 2001, 2006) show that SLA 1996 was effective. Finally, and as we hypothesize in this paper, large data variations and impacts in the later periods might simply overshadow the policy impacts of earlier policy regimes.

### Empirical model and data

For a demonstration purpose, we estimate a monthly econometric model of the U.S. import demand for Canadian lumber using various data sample periods and examine the effects of various policy regimes on U.S. lumber imports. Nagubadi and Zhang (2013) and Parajuli and Zhang (2016) used an import demand function developed by Buongiorno et al. (1979). Following Nagubadi and Zhang (2013) and Parajuli and Zhang (2016), we specify the U.S. imports for Canadian softwood lumber as

\[
q_c = \left\{ \frac{p_{us}, p_{ca}, h, p_{pi}, x_c, \text{MOU}, \text{SLA96}, \text{CVDAD}, \text{SLA06}, \sum_{i=1}^{n} M_i}{H_{11005}} \right\}
\]

where \(q_c\) denotes the monthly U.S. imports from Canadian provinces restricted under the four policy regimes in the month \(t\); \(p_{us}\) and \(p_{ca}\) represent U.S. domestic price and import price of softwood lumber in the U.S., respectively; \(h\) represents the monthly housing starts in the U.S.; \(x_c\) is the real exchange rate between Canadian and U.S. dollars; and \(p_{pi}\) is the overall producer price index for all commodities in the U.S. The policy dummy variables, \(\text{MOU}, \text{SLA96}, \text{CVDAD}, \text{SLA06}\), represent the past trade policies that were imposed in the Canadian lumber shipments to the U.S. (Zhang 2007; Nagubadi and Zhang 2013; Parajuli and Zhang 2016). The dummy terms \(M_i\) are monthly binary variables that capture the seasonal effects in Canadian lumber exports to the U.S. Parajuli and Zhang (2016) explained the expected signs of every variable in detail.

To estimate eq. 1 using the monthly time-series data, we also employ the cointegration framework and vector error correction (VEC) models similar to Nagubadi and Zhang (2013) and Parajuli and Zhang (2016). Because we employ the long time-series data sample from January 1980 to September 2015, the data series might possess potential permanent structural breaks, which might substantially influence the estimation results. Accounting for a structural break in both unit-root test and cointegration analysis, Parajuli and Chang (2015) and Parajuli and Zhang (2016) estimated VEC models for the stumpage market and U.S. import demand for softwood lumber, respectively. In this study, in addition to short-term trend breaks accompanying the trade policy regimes, we consider two permanent structural breaks — January 1987 and January 2006 — in the U.S. softwood lumber market and estimate eq. 1 using the modified cointegration approach of Johansen et al. (2000). The first structural break in January 1987 represents the beginning of trade policy agreements between the U.S. and Canada.
Since then, even during the bilateral trade agreement periods (MOU and SLA06), the trade dispute has been continuously in contention, which carried the dispute case to different international arbitration courts several times (Zhang 2007; Random Lengths 2016). Similarly, the second structural break in January 2006 represents the U.S. housing market bubble in 2006, followed by the great financial crisis in 2008. Equation 1 with two structural breaks is applied to investigate the number of long-run cointegrating vectors in a system of variables. Parajuli and Zhang (2016) used the modified cointegration test in the presence of structural break. If the cointegration test suggests a long-run cointegrating relation among the variables, we can estimate eq. 2 by the VEC method.

To estimate eq. 2 empirically, we employ the same dataset used by Parajuli and Zhang (2016). The historical monthly data for all variables from January 1980 to September 2015 are collected from various sources. Table 1 presents the variables, their descriptions, and their respective data sources. Similar to Parajuli and Zhang (2016), Canadian lumber exports to the U.S. from only SLA-included provinces are taken into account. All price series are deflated to real 1982 dollars using the U.S. producer price index for all commodities. All data series except policy dummy variables and monthly seasonal dummies are log transformed.

### Results

The DF–GLS unit-root test reveals that all variables but \( q_c \) and \( p_p \) are in the integration of order 1(1). Furthermore, the Zivot–Andrews test with an endogenous structural break shows that all variables in eq. 1 are I(1), suggesting that a structural break in \( q_c \) and \( p_p \) affects the power of the DF–GLS unit-root test. Because Parajuli and Zhang (2016) presented unit-root tests, we do not report the results of both unit-root tests here. Table 2 reports the results of the modified cointegration test in the presence of two structural breaks in January 1987 and January 2006. Two-lag vector autoregression specification is selected based on Schwarz’s information criterion. The null hypothesis of no cointegration is rejected, indicating that all six variables in eq. 1 are cointegrated with each other. The test identifies one long-run cointegrating vector at the 1% critical value, suggesting that we can estimate a single equation VEC model of U.S. lumber imports from Canada. We incorporate two trend break deterministic terms, \( \text{td}^{87}_{t-2} \) and \( \text{td}^{06}_{t-2} \) in the system by following the method.
of Joyeux (2007), which capture the effects of the structural break in 1987 and 2006, respectively.

Table 3 presents the long-run and short-run coefficient estimates obtained from the VEC estimation of the U.S. import demand model for Canadian lumber with different data sample periods. The second column of Table 3 reports the estimates of the import model based on the full data sample period of January 1980 to September 2015. The results based on the full data period with two permanent structural breaks show that both MOU and SLA96 are statistically insignificant, suggesting that both past agreements have no significant effects on the Canadian lumber exports to the U.S. These results are consistent with Parajuli and Zhang (2016). However, when we limit our dataset from January 1980 to October 1991 (the last month of MOU), MOU is found to be negative and statistically significant. The short-term VEC estimate in column 3 of Table 3 suggests that, in the period between January 1987 and September 1991, MOU reduces the U.S. lumber imports from Canada by approximately 10%, which is consistent with Wear and Lee (1993).

Similarly, when we subset the dataset from January 1980 to the end of SLA96, the effect of SLA96 is found to be negative and statistically significant, implying that the Canadian lumber shipments to the U.S. were reduced by about 5%. However, the effect of MOU now turns out to be positive and statistically significant, which is similar to the finding of Nagubadi and Zhang (2013) but contrary to the economic theory (column 4 of Table 3). Further, when we extend our data period to September 2006 to incorporate the CVDAD period, SLA96 turns out to be statistically insignificant. The MOU is found to be statistically significant and to remain positive. CVDAD is found to be statistically significant with a coefficient estimate of -0.10 (column 5 of Table 3). It indicates that the period from August 2001 to September 2006, which experienced a widely varying monthly import tariff, caused the U.S. lumber imports from Canada to drop by around 10%.

Thus, it is evident that data sample period selection plays an important role in the empirical analyses of past policy effects. In other words, even if we account for the effects of structural breaks in the long time-series, policy effects in the empirical analysis are time sensitive. As the magnitudes of policy impacts and the market change over time, the larger effects and larger variations in the later period could overshadow the policy impacts in the previous periods.

### Conclusions and discussion

We find that the effects of past trade policies in the U.S. lumber imports model are time sensitive to the ending observation chosen. In the entire data sample period from 1980 to 2015, past trade agreements MOU and SLA96 are found to be statistically insignificant. However, when we subset the data period to the end of each policy regime, effects of both policy regimes are found to be statistically significant with expected signs. We believe that the results using data until the end of each policy regime are consistent with economic theory, empirical results, the purpose of policy analysis, and political economy reality. The latter is demonstrated by the intense fights for the termination of MOU and SLA96. In addition, Fig. 1 clearly shows that MOU stopped the momentum and reversed the trend of increasing Canadian exports of softwood lumber to the U.S. in the late 1980s. Similarly, SLA96 stopped the momentum and possibly reversed the trend of increasing Canadian exports to the U.S. in the middle and late 1990s.

This paper and the studies that we cited are about analysis of policy, not analysis for policy. Analysis of policy often looks at the effect of the current regime or a new policy regime versus a previous policy regime or status quo by comparing current observations with past observations, ceteris paribus. These observations do not need to include future observations because the assumption of all things being equal could be more easily violated with the addition of future observations. As for being consistent with the purpose of policy analysis, we shall use a true story as an example to demonstrate the relationship between the very purpose of policy analysis and the appropriate choice of observations.

A dean of a forestry school evaluates his faculty based on the following five-point scale: 1, “not meet expectation”; 2, “marginally meet expectation”; 3, “meet expectation”; 4, “nearly exceed expectation”; and 5, “exceed expectation”. All faculty in his school have two- or three-way appointments (teaching, research, extension, and service), and for each appointment, he assigns a score of 1–5 as well. The final score for a faculty is the mean, weighted by her individual appointment percentage. If the final score has a
needs consideration. Adding observations after a policy regime is terminated, but not afterwards. Some even advocate a sample size as large as possible. However, qualitative samples size for the asymptotic properties of the estimators. Therefore, it is not fair that a faculty with a score of 4.45 be ranked the same as the faculty with a score of 3.5. The purpose of this analysis of policy exercise was to show that the annual evaluation policy and procedure have flaws and need refinements; two observations were enough.

The dean, on the other hand, provided her with the mean scores of seven other faculty members whose mean score was 4. Statistical analysis shows that 4.45 is not statistically different from the scores of all these faculty at the 5% level. What the dean did was basically add more observations thereby masking the impacts of a fundamental flaw in the annual evaluation system. In doing so, the dean was trying to answer another question: whether the faculty who scored 4.45 performed better than all other faculty whose mean score was 4. This is a different question and irrelevant for the original purpose of the discussion (purpose of policy analysis), which was whether there were flaws in the school’s annual evaluation policy and procedure.

Coming back to our current study, although it is unclear what exactly causes this kind of masking or dilution effects in a long time-series study as we presented in this paper, this is likely related to the facts that impacts in the lumber imports caused by the past policies in the early periods are overshadowed by the larger influences in the later periods (Nagubadi and Zhang 2013), that market changes over time, and that the structure and severity of the trade agreements varies.

In Fig. 1, we see less fluctuation in Canadian lumber imports during the periods of MOU and SLA96 compared with the periods of CVAD and SLA06. Also, the collapse of U.S. housing markets in the recent recession is far longer and more severe than in the early 1980s and in other recessions, whereas the housing markets in the early 2000s were the strongest in the whole study period. As for the structure and severity of the trade agreements, MOU is a 1% export taxes or stumpage adjustment, SLA96 is a two-tier export tax rated quota system, CVAD is a strictly tariff measure although the tariff rate was very high, and SLA06 is a price-specific export tax rated quota system with the possibility of free trade if prevailing monthly U.S. lumber prices reach $355 per thousand board feet. We are not sure which one of these three factors contributes more to this kind of masking or dilution, nor can we answer how much variation in later periods would be required to observe this kind of masking results or if the same sort of dilution happens for estimates in short periods of long data-series regardless of the extent of variations in the later periods. Nonetheless, the larger the dataset that contains observations long after the termination of a policy regime, the harder for one to maintain the assumption of all else being equal and still stick to the very purpose of policy analysis. Thus, we suggest that, in empirical works, policy impacts can best be judged by only using data up to the time when the policy is terminated, but not afterwards.

Obviously, in any statistical analysis, one would need an adequate sample size for the asymptotic properties of the estimators. Some even advocate a sample size as large as possible. However, choosing an appropriate data sample period is quite critical in policy impact analysis studies and the purpose of policy analysis needs consideration. Adding observations after a policy regime is completed does not help assess the impact of the policy regime and often dilutes the impact estimates because it makes harder to maintain the focus of policy analysis and the assumption of all else being equal. Thus, when a series of policy regimes are implemented, a rule for the selection of an appropriate time-series sample in policy analysis should be to include only a specific collection of observations until the end of each policy regime. As for the specific empirical study used in this paper, one could try alternatives to the VEC model such as nonlinear modeling.

Acknowledgements

This work was partially funded through the support of the USDA Forest Service, Southern Research Station, under Joint Venture Agreement # 16-JV-11330143-006. We acknowledge the comments received from two anonymous reviewers and an associate editor of this journal. The usual disclaimer applies.

References


