



University of the
West of England

Faculty of Business and Law

Refining the application of the FLQ Formula for estimating regional input coefficients: an empirical study for South Korean regions

Anthony T. Flegg^a and Timo Tohmo^b

^a University of the West of England, Bristol, UK

^b School of Business and Economics, University of Jyväskylä, Jyväskylä, Finland

*Economics Working Paper Series
1605*



University of the
West of England

bettertogether

Refining the Application of the FLQ Formula for Estimating Regional Input Coefficients: An Empirical Study for South Korean Regions

Anthony T. Flegg

Department of Accounting, Economics and Finance, University of the West of England,
Bristol, UK

E-mail: tony.flegg@uwe.ac.uk

Timo Tohmo

School of Business and Economics, University of Jyväskylä, Jyväskylä, Finland

E-mail: timo.tohmo@jyu.fi

Abstract

This paper uses survey-based data for 16 South Korean regions to refine the application of the FLQ formula for estimating regional input coefficients. Especial attention is paid to the choice of a value for the unknown parameter δ in this formula. Along with regional size, this value determines the size of the adjustment for regional imports in the FLQ formula. Earlier research on this topic using data for two South Korean regions was done by Zhao and Choi (2015). However, using the same basic data, we were unable to replicate their findings. We also identify several methodological shortcomings and some non-trivial computational errors in this pioneering study. We demonstrate that Zhao and Choi have overstated the optimal values of δ for these two regions and understated the FLQ's accuracy. We also establish that the regression model of Kowalewski (2015) is wrongly applied in computing sector-specific values of δ for these two regions. Furthermore, we show that Zhao and Choi's re-estimation of the regression model of Flegg and Tohmo (2013a) yields erroneous results. As well as reworking Zhao and Choi's analysis and extending it from 2 to 16 regions, we make several refinements to Flegg and Tohmo's original model, which was based on data for 20 Finnish regions. Our paper adds to the work of Flegg and Tohmo (2013a, 2016) and Flegg et al. (2016), the underlying aim of which is to find a cost-effective way of adapting national coefficients, so as to produce a satisfactory initial set of regional input coefficients.

1 Introduction

Regional input–output tables are an invaluable aid to regional planning, yet constructing a survey-based regional table can be a complex, expensive and lengthy task. Consequently, regional tables based largely on survey data are rare. A notable exception is South Korea, where the Bank of Korea has constructed survey-based tables for 16 regions for the year 2005. Our main aim is to make full use of this valuable data set to refine the application of the FLQ formula for estimating regional input coefficients. We pay especial attention to the selection of a value for the unknown parameter δ in this formula. Along with regional size, this value determines the size of the adjustment for regional imports in the FLQ formula. Earlier work on this topic using data for two Korean regions was carried out by Zhao and Choi (2015). However, we argue that there are several key shortcomings in this pioneering study, so an effort is made to address these limitations.

In the next section, we discuss the FLQ formula and some related formulae based on location quotients (LQs). The available empirical evidence is also considered. In Section 3,

we present some key findings from Zhao and Choi's study but find that these results cannot be replicated. We then attempt to reconcile our findings with theirs and also raise some fundamental methodological issues pertaining to their approach. Section 4 extends our analysis from 2 to 16 regions, while Section 5 concludes.

2 The FLQ and related formulae

LQs offer a simple and cheap way of regionalizing a national input–output table. In the past, analysts have often used the *simple* LQ (SLQ) or the *cross-industry* LQ (CILQ), yet both are known to underestimate regional trade. This feature is largely attributable to the fact that they either rule out (as with the SLQ) or greatly understate (as with the CILQ) the extent of *cross-hauling* (the simultaneous importing and exporting of a given commodity).¹ The SLQ is defined here as

$$SLQ_i \equiv \frac{Q_i^r / \Sigma_i Q_i^r}{Q_i^n / \Sigma_i Q_i^n} \equiv \frac{Q_i^r}{Q_i^n} \times \frac{\Sigma_i Q_i^n}{\Sigma_i Q_i^r} \quad (1)$$

where Q_i^r is regional output in sector i and Q_i^n is the corresponding national figure. $\Sigma_i Q_i^r$ and $\Sigma_i Q_i^n$ are the respective regional and national totals. Likewise, the CILQ is defined as

$$CILQ_{ij} \equiv \frac{SLQ_i}{SLQ_j} \equiv \frac{Q_i^r / Q_i^n}{Q_j^r / Q_j^n} \quad (2)$$

where the subscripts i and j refer to the supplying and purchasing sectors, respectively.

It should be noted that the SLQ and CILQ are defined in terms of output rather than the more usual employment. Using output is preferable to using a proxy such as employment because output figures are not distorted by differences in productivity across regions. Fortunately, regional sectoral output data were readily available in this instance.

The first step in the application of LQs is to transform the national and regional transactions matrices into matrices of input coefficients. The national coefficient matrix can then be 'regionalized' via the formula

$$r_{ij} = \beta_{ij} \times a_{ij} \quad (3)$$

where r_{ij} is the regional input coefficient, β_{ij} is an adjustment coefficient and a_{ij} is the national input coefficient (Flegg and Tohmo 2016, p. 311). r_{ij} measures the amount of regional input i required to produce one unit of regional gross output j ; it thus excludes any supplies of i obtained from other regions or from abroad. Similarly, a_{ij} excludes any foreign inputs. The role of β_{ij} is to take account of a region's purchases of input i from other regions.

We can estimate the r_{ij} by replacing β_{ij} in equation (3) with an LQ. Thus, for instance:

$$\hat{r}_{ij} = CILQ_{ij} \times a_{ij} \quad (4)$$

No scaling is applied to a_{ij} where $CILQ_{ij} \geq 1$ and likewise for SLQ_i .

The CILQ has the merit that a different scaling can be applied to each cell in a given row of the national coefficient matrix. Unlike the SLQ, the CILQ does not presume that a purchasing sector is either an exporter or an importer of a given commodity but never both.

Even so, empirical evidence indicates that the CILQ still substantially understates regional trade. Flegg et al. (1995) attempted to address this demerit of the CILQ via their FLQ formula, which was later refined by Flegg and Webber (1997). The FLQ is defined here as

$$FLQ_{ij} \equiv CILQ_{ij} \times \lambda^* \quad \text{for } i \neq j \quad (5)$$

$$FLQ_{ij} \equiv SLQ_i \times \lambda^* \quad \text{for } i = j \quad (6)$$

where²

$$\lambda^* \equiv [\log_2(1 + \Sigma_i Q_i^r / \Sigma_i Q_i^n)]^\delta \quad (7)$$

It is assumed that $0 \leq \delta < 1$; as δ increases, so too does the allowance for interregional imports. $\delta = 0$ represents a special case whereby $FLQ_{ij} = CILQ_{ij}$ for $i \neq j$ and $FLQ_{ij} = SLQ_i$ for $i = j$. As with other LQ-based formulae, the constraint $FLQ_{ij} \leq 1$ is imposed.

It is worth emphasizing two aspects of the FLQ formula: its cross-industry foundations and the explicit role given to regional size. With the FLQ, the relative size of the regional purchasing and supplying sectors is considered when making an adjustment for interregional trade. Furthermore, by taking explicit account of a region's relative size, Flegg and Tohmo (2016, p. 312) argue that the FLQ should help to address the problem of cross-hauling, which is apt to be more serious in smaller regions than in larger ones. Smaller regions are likely to be more open to interregional trade.

It is now well established that the FLQ can give more accurate results than the SLQ and CILQ. This evidence includes, for instance, case studies of Scotland (Flegg and Webber 2000), Finland (Tohmo 2004; Flegg and Tohmo 2013a, 2016), Germany (Kowalewski 2015) and Argentina (Flegg et al. 2016). Furthermore, Bonfiglio and Chelli (2008) carried out a Monte Carlo simulation of 400,000 output multipliers. Here the FLQ clearly outperformed its predecessors in terms of yielding the best estimates of multipliers.

A variant of the FLQ is the *augmented* FLQ (AFLQ) formula formulated by Flegg and Webber (2000), which aims to capture the impact of regional specialization on the size of regional input coefficients. This effect is measured via SLQ_j . The AFLQ is defined as

$$AFLQ_{ij} \equiv FLQ_{ij} \times \log_2(1 + SLQ_j) \quad (8)$$

The specialization term, $\log_2(1 + SLQ_j)$, only applies where $SLQ_j > 1$ (Flegg and Webber 2000, p. 566). The AFLQ has the novel property that it can encompass cases where $r_{ij} > a_{ij}$ in equation (3). As with the FLQ, the constraint $AFLQ_{ij} \leq 1$ is imposed.

Nonetheless, although the AFLQ has some theoretical merits relative to the FLQ, its empirical performance is typically very similar. For instance, in the Monte Carlo study by Bonfiglio and Chelli (2008), the AFLQ gave only slightly more accurate results than the FLQ.³ This outcome was confirmed by Flegg et al. (2016). Kowalewski (2015) also tested both formulae but again obtained similar results, as did Zhao and Choi (2015, tables 7 and 8). For this reason, along with limitations of space, only the FLQ will be examined here.

Another variant of the FLQ is proposed by Kowalewski (2015). Her innovative approach involves relaxing the assumption that the value of δ is invariant across regions. The effectiveness of this approach is assessed by Zhao and Choi (2015), whose findings are re-examined in the next section.

Kowalewski's industry-specific FLQ, the SFLQ, is defined as

$$SFLQ_{ij} \equiv CILQ_{ij} \times [\log_2(1 + E^r/E^n)]^{\delta_j} \quad (9)$$

where E^r/E^n is regional size measured in terms of employment. In order to estimate the values of δ_j , Kowalewski specifies a regression model of the following form

$$\delta_j = \alpha + \beta_1 CL_j + \beta_2 SLQ_j + \beta_3 IM_j + \beta_4 VA_j + \varepsilon_j \quad (10)$$

where CL_j is the coefficient of localization, which measures the degree of concentration of national industry j , IM_j is the share of foreign imports in total national intermediate inputs, VA_j is the share of value added in total national output and ε_j is an error term. Regional data are needed for SLQ_j , whereas CL_j , IM_j and VA_j require national data. CL_j is calculated as

$$CL_j \equiv 0.5 \sum_r \left| \frac{E_j^r}{E_j^n} - \frac{E^r}{E^n} \right| \quad (11)$$

The FLQ's focus is on the output and employment generated within a specific region. As Flegg and Tohmo (2013b) point out, it should only be used in conjunction with national input–output tables where the inter-industry transactions exclude imports (type B tables). By contrast, where the focus is on the overall supply of goods, Kronenberg's Cross-Hauling Adjusted Regionalization Method (CHARM) can be employed (Flegg et al. 2015; Többen and Kronenberg 2015). CHARM requires type A tables, those where imports have been incorporated into the national transactions table (Kronenberg 2009, 2012).

Nevertheless, Zhao and Choi (2015) employ type A national tables in their study. This choice can be expected to cause an upward bias in the estimated type I output multipliers. Furthermore, the optimal values of δ are likely to be overstated and the accuracy of the FLQ understated. These expected outcomes are confirmed in the next section.

3 Zhao and Choi's study

3.1 The SLQ, CILQ and FLQ

Zhao and Choi based their analysis on a 28×28 national technological coefficient matrix for 2005 produced by the Bank of Korea. The authors regionalized this survey-based matrix by applying various LQ-based formulae calculated using employment data. The Bank divided the country into 16 regions and computed type I output multipliers for each region. Zhao and Choi chose to study two regions in detail, namely Daegu and Gyeongbuk, and used the Bank's regional multipliers as a benchmark for assessing the accuracy of their simulations. As criteria, they used the mean absolute distance and the mean absolute percentage error (MAPE). However, the results from these two measures hardly differed, so only MAPE will be considered here. It was calculated via the formula

$$MAPE = (100/n) \sum_j |\hat{m}_j - m_j| / m_j \quad (12)$$

where m_j is the type I output multiplier for sector j and $n = 28$ is the number of sectors.

Table 1 near here

A selection of Zhao and Choi's results is presented in Table 1. As expected, the FLQ outperforms the SLQ and CILQ, yet the extent of this superior performance is striking. It echoes the clear-cut findings in the Monte Carlo study of Bonfiglio and Chelli (2008), yet other authors such as Flegg and Tohmo (2013a, 2016), Flegg et al. (2016) and Kowalewski

(2015) have found more modest differences in performance. An interesting facet of the results is that MAPE is minimized at a relatively high value of δ in both regions. However, most other studies, including those mentioned above, have found much lower optimal values.

Table 2 near here

As the first step in our evaluation of Zhao and Choi's study, we attempted to replicate their results using the assumptions stated in their paper. Our findings are displayed in Table 2. To attain greater accuracy, we used steps of 0.05 for δ . It is evident that the results are somewhat different. Having checked our own calculations carefully, it would seem that errors of an unknown nature must have occurred in Zhao and Choi's simulations.⁴ Daegu is most affected by the re-estimation, with a cut in the optimal δ from 0.5 to 0.4. The performance of the SLQ and CILQ is also noticeably better. By contrast, for Gyeongbuk, the optimal δ is still 0.6, although the corresponding value of MAPE has risen from 8.5% to 10.5%. The values of MAPE for the SLQ and CILQ in this region are little changed by the re-estimation.

As noted earlier, a problem with Zhao and Choi's approach is their use of a type A national coefficient matrix, which would have the effect of overstating the optimal values of δ . The explanation is straightforward: instead of using the equation $\hat{r}_{ij} = FLQ_{ij} \times a_{ij}$ to estimate the input coefficients, one would be using the equation $\hat{r}_{ij} = FLQ_{ij} \times (a_{ij} + f_{ij})$, where f_{ij} is the national propensity to import from abroad. Minimizing MAPE would then require a higher value of δ .⁵

Table 3 near here

Table 3 illustrates the consequences of using a type A rather than type B national coefficient matrix. The most striking change compared with Table 2 occurs in Gyeongbuk: there is a big fall in the optimal δ from 0.6 to 0.35, while the corresponding value of MAPE is cut from 10.5% to 6.5%. For Daegu, the optimal δ remains at 0.4 but MAPE is lowered from 9.2% to 7.1%. There is a further improvement in the performance of the SLQ and CILQ, although these conventional LQs are still substantially less accurate than the FLQ.

Table 4 near here

It was noted earlier that, wherever possible, it is preferable to use output rather than employment data since interregional differences in productivity are liable to distort any calculations based on employment. Table 4 records the effects of using sectoral output data to compute the SLQ, CILQ and FLQ. One can see that the optimal values of δ are cut by 0.05 in both regions and there is a modest fall in the corresponding MAPEs. Once more, an enhanced performance of the SLQ and CILQ is evident in both regions.

Table 5 near here

The consequences of switching from an employment-based to an output-based measure of regional size are captured in Table 5. This change only affects the FLQ since the SLQ and CILQ do not incorporate an adjustment for regional size. In terms of employment, regional size is 4.7% for Daegu and 5.4% for Gyeongbuk; the corresponding figures for output are 2.9% and 8.4%, respectively. As expected, Table 5 reveals that using output rather than employment to measure regional size has the effect of lowering the required δ for Daegu, yet raising it for Gyeongbuk. In effect, Daegu has become a smaller region, while Gyeongbuk

has become a larger one. Therefore, the FLQ would make a bigger allowance for regional imports in Daegu but a smaller one in Gyeongbuk. The required values of δ would change as a result. There is, however, negligible impact on the corresponding values of MAPE.

From the above discussion, it seems fair to conclude that Zhao and Choi (2015) have substantially overstated the required values of δ and understated the accuracy of the FLQ. Also, even though the FLQ is still demonstrably more accurate than the SLQ and CILQ, the extent of this superiority is less marked than their results would suggest.

3.2 The sector-specific approach using the SFLQ

A key part of Zhao and Choi's study is a test of a new sector-specific FLQ formula, the SFLQ, devised by Kowalewski (2015). As explained earlier, this method involves using the regression model (10) to generate sector-specific values of δ for each region. Kowalewski's results for a German region are reproduced in Table 6, along with Zhao and Choi's Korean findings and our own estimates. For consistency, we recomputed the SLQ_j using sectoral employment data.

Table 6 near here

Looking first at Kowalewski's results, it is striking how one of the regressors, CL_j , is highly statistically significant, whereas the remaining three have low t statistics. The positive estimated coefficient of CL_j is consistent with Kowalewski's argument that 'the more an industry is concentrated in space, the higher the regional propensity to import goods or services of this industry' (Kowalewski 2015, p. 248). Such industries would require a higher value of δ to adjust for this higher propensity. As expected, SLQ_j has a negative estimated coefficient. Kowalewski's rationale here is that 'regional specialization would lead to an increase in intra-regional trade and a decrease in imports', so that 'one would expect a higher SLQ_j to be accompanied by a lower value of δ_j , which would additionally (to the FLQ formula) dampen regional imports' (Kowalewski 2015, p. 248). However, the t statistic for SLQ_j is very low, which suggests that this variable may not be relevant. Likewise, the results for both IM_j and VA_j cast doubt on their relevance.

Zhao and Choi's results are puzzling. Kowalewski's method requires a separate regression for each region since the values of SLQ_j would vary across regions. However, the authors report results for only one regression and offer no explanation as to how it was estimated or to which region it relates. Moreover, the estimated coefficient of CL_j is implausibly large and is markedly out of line with both Kowalewski's estimate and our own figures for Daegu and Gyeongbuk. The credibility of Zhao and Choi's results is also undermined by the fact that they were derived from a type A national coefficient matrix.

Turning now to our own regressions, the results for Daegu look sensible on the whole. The R^2 is only a little below that reported by Kowalewski. Moreover, CL_j is statistically significant at the 1% level and its estimated coefficient has the anticipated sign. Although SLQ_j and VA_j are still not significant at conventional levels, their t ratios are much better than in Kowalewski's regression. As for IM_j , this does indeed seem to be a redundant variable.

Our regression for Gyeongbuk leaves much to be desired in terms of both goodness of fit and the outcomes for CL_j and SLQ_j . However, a redeeming feature is the highly statistically significant result for VA_j . Kowalewski does not offer a rationale for including this variable but one might argue that a higher share of value added in total national output would mean a lower share of intermediate inputs and hence lower imports. If this effect were transmitted to regions, it is possible that a lower δ_j would be needed, i.e. $\beta_4 < 0$ in equation (10).

Table 7 near here

Table 7 displays estimates of δ_j derived from our regressions, along with the ‘optimal’ values that would minimize MAPE for the type I output multipliers. To compute the optimal δ_j , we performed the calculations on a sectoral basis, using steps of 0.025 for δ , and then applied linear interpolation.

To shed some light on the quality of our estimates, we correlated $\hat{\delta}_j$ with δ_j . The simple correlation coefficient, r , was 0.739 ($p = 0.000$) for Daegu and 0.640 ($p = 0.000$) for Gyeongbuk. The fact that both correlations are highly statistically significant lends support to Kowalewski’s approach, although there is clearly still much scope for enhanced accuracy. As regards the difference in the size of r , this reflects the fact that Table 6 shows a higher R^2 for Daegu than for Gyeongbuk.

In commenting on their results, Zhao and Choi (2015, pp. 911–913) highlight two contrasting sectors, namely ‘Petroleum and coal products’ (7) and ‘Finance and insurance’ (23), which they note are the smallest and largest industries, respectively. On the basis of their findings that δ_{23} equalled 0.1 for Daegu and 0.2 for Gyeongbuk, whereas $\delta_7 = 0.9$ for both regions, they posit an inverse relationship between a sector’s size and the value of δ_j . Whilst this hypothesis is plausible, we were unable to replicate their supporting calculations. Indeed, instead of $\delta_7 = 0.9$ for both regions, Table 7 gives values of 0.000 for Daegu and 0.369 for Gyeongbuk.⁶ As for δ_{23} , our figures are 0.049 and 0.035, respectively. More generally, there is a far from perfect association between their set of optimal values and ours: $r = 0.548$ ($p = 0.003$) for Daegu and 0.573 ($p = 0.002$) for Gyeongbuk.

Table 8 near here

The relative performance of the SFLQ in terms of MAPE is examined in Table 8. The table distinguishes between cases where the δ_j have been estimated via regressions and those where optimal values have been used. Of course, in reality, analysts using non-survey methods would not know the optimal values, so the results illustrate the best outcomes that could be attained with the SFLQ in a perfect world. Based on our calculations, a residual error of about 2% would remain in each region.

Table 8 also reveals that our SFLQ regressions yield a lower MAPE than the FLQ in Daegu but less obviously so in Gyeongbuk. For instance, in the absence of any other information, $\delta = 0.3$ is a value an analyst might use for the FLQ. On that basis, the potential gain from using the SFLQ, in conjunction with our regressions, would be 1.5 percentage points in Daegu but only 0.3 in Gyeongbuk.

In discussing their findings, Zhao and Choi (2015, p. 913) comment that it is ‘undeniable that SFLQ presents an extraordinary ability to minimize errors produced by regionalization’. However, this statement is based on a comparison with results derived using optimal values. We would argue that the only relevant comparison is with regression-based estimates, which would be the only information potentially available to an analyst using non-survey data. Clearly, with a MAPE of 19.5 for Daegu and 15.7 for Gyeongbuk, Zhao and Choi’s results would not be helpful in that respect.

It is evident that Kowalewski’s SFLQ approach can yield a useful, albeit modest, enhancement of accuracy relative to the FLQ if used in conjunction with a well-specified regression model. Zhao and Choi (2015, p. 915) suggest that possible ways of refining these regressions could include (i) introducing new explanatory variables and (ii) using non-linear formulations. Unfortunately, it is hard to think of new variables for which data would be readily available. As regards refinement (ii), we considered the following alternative non-

linear models:

$$\ln \delta_j = a + b_1 CL_j + b_2 SLQ_j + b_3 IM_j + b_4 VA_j + e_j \quad (13)$$

$$\ln \delta_j = c + d_1 \ln CL_j + d_2 \ln SLQ_j + d_3 \ln IM_j + d_4 \ln VA_j + f_j \quad (14)$$

Table 9 near here

Table 9 reports a mixed outcome: the linear model is best for Daegu, whereas the double-log model (14) is best for Gyeongbuk. However, the differences in performance of the three models are not substantial.

4 Extension to all regions

4.1 Results for 16 regions

In this section, we expand our analysis to encompass all 16 South Korean regions. The main objective is to produce results that are more generally valid, particularly in terms of finding appropriate values for the unknown parameter δ in the FLQ formula. Before considering our findings, it may be helpful to examine the key regional characteristics presented in Table 10.

Table 10 near here

Table 10 examines two alternative ways of measuring regional size. Although one can see at a glance that the output and employment shares are not perfectly matched, there is agreement that Gyeonggi-do and Seoul are the two biggest regions and that Jeju-do is the smallest. Clearly, with $r = 0.921$, one would not expect the choice of measure to make a major difference to most outcomes for the FLQ. Even so, output would be our preferred measure since it is not distorted by interregional variations in productivity.

LQ-based approaches presuppose that the share of foreign imports in gross output does not vary across regions. Any deviations from the mean would introduce inaccuracies into the simulations, yet Table 10 reveals that this assumption of constant shares is far from being satisfied. Ulsan, in particular, stands out as having an especially high share of foreign imports. As regards imports from other regions, Seoul and Jeollanam-do exhibit noticeably lower shares than those in the other regions, where the shares do not vary much from the mean. In fact, the coefficient of variation, V , shows that the dispersion in this variable is low. Accounting for differences in the propensity to purchase inputs from other regions is, of course, something that the FLQ is designed to do.

Herfindahl's index, $H_r = \sum_i (Q_i^r / \sum_i Q_i^r)^2$, where Q_i^r is the output of sector i in region r , measures the extent to which each region's output is concentrated in one or more sectors. Ulsan again stands out as having an unusually high value for H_r . However, apart from Seoul, Gyeongbuk and Jeollanam-do, the remaining regions all have values fairly close to the mean.

Table 11 near here

The minimum MAPE in each region is identified in bold in Table 11, along with the corresponding optimal value of δ .⁷ It is evident that there is much interregional variation in the optimal values of δ , yet it is also true that ten of these values lie in the range 0.4 ± 0.05 , where MAPE is about 8%. Gangwon-do and Jeju-do are atypical in requiring $\delta = 0.2$. On the other hand, three regions need at least $\delta = 0.5$.⁸ Looking at the overall pattern of results, there

does seem to be some tendency for the minimum MAPE to rise along with the optimal δ .

4.2 Sensitivity analysis using different criteria

The simulations thus far have been evaluated solely in terms of MAPE, thereby facilitating comparisons with the work of Zhao and Choi (2015). Although MAPE has some desirable properties as a criterion, it does not capture all aspects relevant to the choice of method. It is desirable, therefore, to employ a range of criteria with different properties. Following previous research (Flegg and Tohmo 2013a, 2016; Flegg et al. 2016), the following additional statistics will be employed to assess the accuracy of the estimated multipliers:

$$\text{MPE} = (100/28) \sum_j (\hat{m}_j - m_j) / m_j \quad (15)$$

$$\text{WMPE} = 100 \sum_j w_j (\hat{m}_j - m_j) / m_j \quad (16)$$

$$\text{SDSD} = \{\text{sd}(\hat{m}_j) - \text{sd}(m_j)\}^2 \quad (17)$$

$$U = 100 \sqrt{\frac{\sum_j (\hat{m}_j - m_j)^2}{\sum_j m_j^2}} \quad (18)$$

MPE is the mean percentage error. This statistic has been added to the set of criteria because it offers a convenient way of measuring the amount of bias in a relative sense. It has also been used in many previous studies. WMAE is the weighted mean percentage error, which takes into account the relative importance of each sector. w_j is the proportion of total regional output produced in sector j . The role of the squared difference in standard deviations (SDSD) is to assess how far each method is able to replicate the dispersion of the benchmark distribution of multipliers. Finally, U is Theil's well-known index of inequality, which has the merit that it encompasses both bias and variance (Theil et al. 1966, pp. 15–43). A selection of results is presented in Table 12.

Table 12 near here

Table 12 reveals a high degree of consistency in the results across different criteria. In particular, regardless of which criterion is used, the SLQ and CILQ yield very similar results and both perform very poorly indeed relative to the FLQ. The MPE shows, for example, that the SLQ would overstate the sectoral multipliers by 21.2% on average across the 16 regions, whereas the FLQ with $\delta = 0.35$ would reduce this overstatement to 1.2%. With a δ slightly above 0.375, the bias would vanish. What is more, apart from the sign, the MPE and MAPE for the SLQ are very similar, which shows that the simulation errors are nearly always overestimates. This is true too for the CILQ.

The fact that MPE, SDSD and U all indicate an optimal $\delta \approx 0.375$ is interesting as it shows that there is no conflict between minimizing bias and variance in this data set. It is also evident that the accuracy of the FLQ greatly surpasses that of the SLQ and CILQ in both respects. Furthermore, one would not go far wrong in setting $\delta = 0.375$, although it is true that WMPE indicates $\delta = 0.3$. This difference suggests that a somewhat smaller δ may be required for the relatively larger sectors. It is worth noting that a δ in the range 0.4 ± 0.025 would yield MAPE $\approx 8.0\%$.

4.3 Choosing values for δ

Although the results presented earlier offer some guidance regarding appropriate values of δ , it would be helpful if a suitable estimating equation could be developed. With this aim in mind, Flegg and Tohmo (2013a, p. 713) fitted the following model to survey-based data for twenty Finnish regions in 1995:

$$\ln \delta = -1.8379 + 0.33195 \ln R + 1.5834 \ln P - 2.8812 \ln I + e \quad (19)$$

where R is regional size measured in terms of output and expressed as a percentage; P is the proportion of each region's gross output imported from other regions, averaged over all sectors and divided by the mean for all regions; I is each region's average use of intermediate inputs (including inputs from other regions), divided by the corresponding national average; e is a residual. Observations on $\ln \delta$ were derived by finding the value of δ that minimized MPE for each Finnish region. $R^2 = 0.915$ and the t ratios for the three regressors were 11.66, 6.25 and -3.33 , respectively. The model comfortably passed the χ^2 diagnostic tests for functional form, normality and heteroscedasticity.

Table 13 near here

Table 13 records the results of our re-estimation of Flegg and Tohmo's model using data for all 16 South Korean regions.⁹ Observations on $\ln \delta$ were derived by finding the value of δ that minimized MAPE for each region.¹⁰ Regression (1) has the same specification as equation (19) and the corresponding estimated elasticities have identical signs. However, when judged in terms of the usual statistical criteria, this new model is less satisfactory than the Finnish one. As a result, we attempted to refine it by adding a new regressor, $\ln F$, where F is the average proportion of each region's gross output imported from abroad, divided by the mean for all regions. As noted earlier and illustrated in Table 10, the share of foreign imports in gross output varies greatly across regions, so this variable should be relevant.

It is evident that $\ln F$ adds greatly to the explanatory power of the model and its estimated coefficient has the anticipated sign. However, the χ^2 statistic reveals that the residuals are not normally distributed. Daejeon was identified as the main source of this problem: its residual is more than two standard errors from zero. To address this problem, and to prevent this outlier from distorting the results, a binary variable, B_{15} , was added to the model.¹¹ Regression (3) records the outcome.

It is clear from the χ^2 statistic that the inclusion of B_{15} has gone a long way towards eliminating the skewness and kurtosis in the residuals of regression (2). The big rise in R^2 reflects the fact that B_{15} is highly statistically significant. There is also a marked rise in the t ratios for $\ln R$, $\ln P$ and $\ln F$. However, the results strongly suggest that $\ln I$ is a redundant regressor, so it has been omitted from regression (4).

There are several reasons why it is desirable to omit $\ln I$: to simplify the model, to save a valuable degree of freedom and to reduce unnecessary multicollinearity. Despite the theoretical reasons for including $\ln I$, its negligible t ratio reflects the fact that there is relatively little interregional variation in the values of $\ln I$. The benefits of omitting $\ln I$ are demonstrated by the fact that every t ratio has improved. Moreover, there is minimal change in the estimated coefficients, which indicates that omitting $\ln I$ has not introduced bias. One can see too that regression (4) has the highest AIC, thereby confirming that it has the best fit.¹² It also comfortably passes all χ^2 diagnostic tests.

Regression (4) differs in some key respects from the Finnish equation (19). In particular, $\ln F$ plays a highly significant role in regression (4), yet is absent from equation (19), whereas

$\ln I$ is highly significant in equation (19) but does not appear in regression (4). As regards regional size, the estimated elasticity of δ with respect to R is 0.332 in equation (19), yet only 0.168 in regression (4). These dissimilarities can largely be explained by the differences between Finland and South Korea in the amount of interregional variation in each variable. Although we tried to refine the regressions by adding $\ln H$, where H is Herfindahl's index of concentration, $\ln H$ had a negligible t ratio. The explanation is probably that H does not vary much across South Korean regions (see Table 10). Flegg and Tohmo (2013a, note 26) report a similar outcome for Finland.

Before assessing how accurately regression (4) is able to estimate δ for individual regions, it is worth examining an alternative approach proposed by Bonfiglio (2009), who used simulated data from a Monte Carlo study to derive the following regression equation:

$$\hat{\delta} = 0.994 \text{ PROP} - 2.819 \text{ RSRP} \quad (20)$$

where PROP is the propensity to interregional trade (the proportion of a region's total intermediate inputs bought from other regions) and RSRP is the relative size of regional purchases (the ratio of total regional to total national intermediate inputs). The principal advantage of a Monte Carlo approach, according to Flegg and Tohmo (2016, p. 33), lies in the generality of the findings, whereas 'the results derived from a single region may reflect the peculiarities of that region and thus not be valid in general.' Even so, Flegg and Tohmo (2016, p. 33) remark that 'the simplifying assumptions underlying a Monte Carlo simulation mean that it cannot replicate the detailed economic structure and sectoral interrelationships of regional economies.'¹³ Also, with data for 16 regions, concerns about lack of generality are less compelling here, although there remains the possibility that South Korea is a unique case.

Table 14 near here

The first column in Table 14 displays the optimal values of δ , those that minimize MAPE for the sectoral multipliers, while the second column records the predicted values from regression (4) in Table 13. There is a very close correspondence between the two sets of values, with $r = 0.957$ ($p = 0.000$). This outcome reflects the high R^2 of regression (4). By contrast, Bonfiglio's method gives very poor estimates of δ and there is a negative, rather than positive, correlation between $\hat{\delta}$ and δ , with $r = -0.485$ ($p = 0.057$).¹⁴ Moreover, $\hat{\delta} < 0$ for the two largest regions, which contradicts the theoretical restriction that $\delta \geq 0$. Flegg and Tohmo (2016, p. 33) note that $\hat{\delta} < 0$ can occur where regions are relatively large or exhibit below-average propensities to import from other regions or have both characteristics. They identify two such regions. In view of these problems, we would not recommend the use of Bonfiglio's method.¹⁵

With respect to Flegg and Tohmo's approach, the form in which regression (4) in Table 13 is specified should make it simpler for analysts to derive a figure for δ . The equation, with $B_{15} = 0$, is reproduced below.

$$\ln \delta = -1.2263 + 0.1680 \ln R + 0.3254 \ln P + 0.3170 \ln F + e \quad (19)$$

An analyst using this equation would need to make an informed assumption about how far a region's propensity to import from other regions diverged from the mean for all regions in a country, which should be easier than having to measure this propensity directly. Likewise, an allowance could be made for any assumed divergence between the regional and national shares of foreign inputs. It would also be easy to carry out a sensitivity analysis. However,

in some cases, it might be more convenient for an analyst to work with the following alternative version of equation (19):

$$\ln \delta = -3.0665 + 0.1680 \ln R + 0.3254 \ln p + 0.3170 \ln f + e \quad (20)$$

where p is each region's propensity to import from other regions and f is each region's average use of foreign intermediate inputs, both measured as a proportion of gross output.¹⁶

To assess the accuracy of equation (19) in terms of sectoral multipliers, again consider Table 14. For each region, the specified value of δ was used in estimating the sectoral multipliers and hence MAPE for that region. The results were then averaged over all regions to get MAPE $\approx 7.3\%$. By contrast, Table 12 reveals that using a common $\delta = 0.4$ for all regions would give MAPE $\approx 8.0\%$, so there would be a gain of about 0.7 percentage points, on average, from using the region-specific equation (19).

5 Conclusion

This paper has employed survey-based data for 16 South Korean regions to refine the application of the FLQ formula for estimating regional input coefficients. The focus was on the choice of a value for the unknown parameter δ in this formula. Along with regional size, this value determines the size of the adjustment for regional imports in the FLQ formula.

At the outset, we re-examined the pioneering study of Zhao and Choi (2015), who explored this topic using data for two South Korean regions, Daegu and Gyeongbuk. However, using the same basic data, we were unable to replicate their findings. We also identified several methodological shortcomings and some non-trivial computational errors. The shortcomings included the use of a national coefficient matrix that included imports. As a result, we found that Zhao and Choi had overstated the optimal values of δ for these two regions and understated the FLQ's accuracy. We also demonstrated that the regression models of both Flegg and Tohmo (2013a) and Kowalewski (2015) had not been applied correctly and consequently did not yield valid results. As well as reworking Zhao and Choi's analysis and extending it from 2 to 16 regions, we made several refinements to Flegg and Tohmo's original model, which was based on data for 20 Finnish regions, and carried out a rigorous test of Kowalewski's sector-specific approach.

Several important findings emerged from our statistical analysis. For instance, on average across the 16 regions, the FLQ with δ in the range 0.4 ± 0.025 gave a minimum mean absolute percentage error (MAPE) $\approx 8.0\%$ when estimating the type I sectoral output multipliers, compared with 23.5% for the CILQ and 22.2% for the SLQ. The mean percentage error for the FLQ with $\delta = 0.375$ was -0.1% , as against -22.4% for the CILQ and -21.2% for the SLQ. Although it is unsurprising that the CILQ and SLQ should yield overstated multipliers, the size of this bias is striking. The credibility of these findings is bolstered by the fact that they were confirmed by Theil's inequality coefficient, which takes both bias and dispersion into account.

We gave detailed consideration to the sector-specific approach proposed by Kowalewski (2015), which aims to enhance the accuracy of the FLQ by permitting δ to vary across sectors. Her SFLQ approach involves using a regression model to estimate a value of δ for each sector j in a given region. We re-estimated this model for the two regions studied by Zhao and Choi, using as regressors a region-specific variable, SLQ_j , and three other variables based on national data.

However, we identified only modest gains from using the SFLQ: for example, relative to the FLQ with $\delta = 0.3$, MAPE was reduced by 1.5 and 0.3 percentage points in Daegu and Gyeongbuk, respectively. As the accuracy of the SFLQ depends crucially on the regression

model used to estimate the δ_j , more research is clearly needed to improve its specification.

Unlike the SFLQ, Flegg and Tohmo's model uses region-specific data exclusively. In our reformulation of this model using South Korean data, we included regressors to capture regional size and the propensities to import from other regions and from abroad. Interregional variation in the propensity to import from abroad played a key role in determining the value of δ . The model was found to be satisfactory in terms of a range of statistical criteria and gave relatively accurate estimates of δ . Using this model to derive region-specific estimates of δ lowered MAPE by some 0.7 percentage points on average.

It would be fair to conclude that the findings in this paper offer support for employing the FLQ as a non-survey regionalization technique. Nonetheless, as with other pure non-survey methods, it can only be relied upon to produce a satisfactory initial set of regional input coefficients. Analysts should always endeavour to refine these estimates via informed judgement, making use of any available superior data, carrying out surveys of key sectors and so on. Indeed, we would argue that the FLQ is very well suited to building the non-survey foundations of a hybrid model.¹⁷ Consideration should also be given to the use of Flegg and Tohmo's regression model, as reformulated in this study, and to the SFLQ. It is worth noting, finally, that interesting recent work by Hermannsson (2016) and Jahn (2016) has extended the use of the FLQ from an analysis of single regions to a multi-regional context.

Footnotes

1. See Flegg and Tohmo (2013b, p. 239 and note 3).
2. Cf. Flegg and Webber (1997, p. 798), who define λ^* in terms of employment. This reflects the fact that, in most cases, employment has to be used as a proxy for output.
3. For instance, the minimum mean relative absolute distance was 19.1% for the FLQ (with $\delta = 0.3$) but 18.3% for the AFLQ (with $\delta = 0.4$). See Bonfiglio and Chelli (2008, table 1).
4. We are grateful to Professors Zhao and Choi for letting us examine their data. This enabled us to verify that we were using the same sectoral classifications, national transactions matrix, employment data and LQs. Even so, we were still unable to replicate their findings.
5. Zhao and Choi (2015, footnote 2) recognize that the use of a type A national table would yield overestimated regional multipliers, yet they assert that this is immaterial if one's aim is to ascertain which technique is the best for generating regional tables. However, even if true, this is a very limited view of the purpose of detailed empirical studies in this area, which are needed in order to measure how far the performance of alternative methods differs and, crucially, to offer guidance regarding appropriate values of δ . Zhao and Choi's study does not yield reliable information on those aspects.
6. Actually, the $\delta_7 = 0.000$ shown in Table 7 for Daegu is a unique case where the minimum MAPE occurred at a negative value of δ ; this is theoretically unacceptable, so a zero value was imposed. When calculating optimal values, Zhao and Choi used steps of 0.1 for δ_j rather than 0.025, and did not interpolate between values.
7. The results for Daegu and Gyeongbuk in Table 11 are a little different from those in Table 5. This is because we used our own calculations of benchmark multipliers for Table 11 but the Bank of Korea's figures for Table 5, to ensure consistency with Zhao and Choi's data.
8. The optimal δ is approximately 0.534 for Gyeonggi-do, 0.542 for Ulsan and 0.497 for Daejeon.
9. Zhao and Choi (2015, table 2) report the results of estimating, using South Korean data, what they refer to as 'Flegg's model'. However, this regression has an $R^2 = 0.003$ and regional size, R , is the sole explanatory variable. It is not explained how this result was obtained. By contrast, when we regressed $\ln \delta$ on $\ln R$ alone, the $R^2 = 0.394$.
10. To estimate a value yielding the minimum MAPE in each region, we varied δ in steps of 0.0001.
11. $B_{15} = 1$ for Daejeon and zero otherwise. As the second smallest region, Daejeon is atypical in the sense that it requires an unusually high value of $\delta \approx 0.5$. Without B_{15} , $\hat{\delta} = 0.306$ for this region.
12. $AIC = \ln L - (k + 1)$, where $\ln L$ is the maximized log-likelihood of the regression and k is the number of regressors. Compared with the more conventional \bar{R}^2 , AIC takes more account of k .

13. For instance, Bonfiglio and Chelli (2008, p. 248) generated their regional input and import coefficients randomly in the interval $[0, 1]$, yet that range does not represent a realistic representation of a real regional table, where input coefficients tend to be small, except for those along the principal diagonal.
14. We used output shares (see Table 10) to proxy RSRP. For PROP, we used the ratio A/B , where A = imports from other South Korean regions, and B = A + intraregional intermediate inputs + imports from abroad.
15. For a more detailed evaluation of Bonfiglio's method, see Flegg and Tohmo (2016, pp. 33–34).
16. See Flegg and Tohmo (2016, pp. 34–35) for a more detailed discussion of this approach.
17. See Jackson (1998) and Lahr (1993, 2001) concerning the hybrid approach.

Acknowledgements

We are grateful to Professors Zhao and Choi for letting us examine their data. Jihyeon Lee of the Bank of Korea helped us to get started on this investigation. We would also like to thank Jeroen Klijs and Chris Webber for their helpful comments.

References

- Bonfiglio, A. (2009) On the parameterization of techniques for representing regional economic structures. *Economic Systems Research*, 21, 115–127.
- Bonfiglio, A. and Chelli, F. (2008) Assessing the behaviour of non-survey methods for constructing regional input–output tables through a Monte Carlo simulation. *Economic Systems Research*, 20, 243–258.
- Flegg, A.T. and Tohmo, T. (2013a) Regional input–output tables and the FLQ formula: a case study of Finland. *Regional Studies*, 47, 703–721.
- Flegg, A.T. and Tohmo, T. (2013b) A comment on Tobias Kronenberg's 'Construction of regional input–output tables using nonsurvey methods: the role of cross-hauling'. *International Regional Science Review*, 36, 235–257.
- Flegg, A.T. and Tohmo, T. (2016) Estimating regional input coefficients and multipliers: the use of FLQ is not a gamble. *Regional Studies*, 50, 310–325.
- Flegg, A.T., Huang, Y. and Tohmo, T. (2015) Using CHARM to adjust for cross-hauling: the case of the province of Hubei, China. *Economic Systems Research*, 27, 391–413.
- Flegg, A.T., Mastronardi, L.J. and Romero, C.A. (2016) Evaluating the FLQ and AFLQ formulae for estimating regional input coefficients: empirical evidence for the province of Córdoba, Argentina. *Economic Systems Research*, 28, 21–37.
- Flegg, A.T. and C.D. Webber (1997) On the appropriate use of location quotients in generating regional input–output tables: reply. *Regional Studies*, 31, 795–805.
- Flegg, A.T. and Webber, C.D. (2000) Regional size, regional specialization and the FLQ formula. *Regional Studies*, 34, 563–569.
- Flegg, A.T., Webber, C.D. and Elliott, M.V. (1995) On the appropriate use of location quotients in generating regional input–output tables. *Regional Studies*, 29, 547–561.
- Hermannsson, K. (2016) Beyond intermediates: the role of consumption and commuting in the construction of local input–output tables. *Spatial Economic Analysis*, forthcoming.
- Jackson, R.W. (1998) Regionalizing national commodity-by-industry accounts. *Economic Systems Research*, 10, 223–238.
- Jahn, M. (2016) Extending the FLQ formula: a location quotient-based interregional input–output framework. *Regional Studies*, forthcoming.
- Kowalewski, J. (2015) Regionalization of national input–output tables: empirical evidence on the use of the FLQ formula. *Regional Studies*, 49, 240–250.
- Kronenberg, T. (2009) Construction of regional input–output tables using nonsurvey methods: the role of cross-hauling. *International Regional Science Review*, 32, 40–64.
- Kronenberg, T. (2012) Regional input–output models and the treatment of imports in the European system of accounts. *Jahrbuch für Regionalwissenschaft (Review of Regional Research)*, 32, 175–191.

- Lahr, M.L. (1993). A review of the literature supporting the hybrid approach to constructing regional input–output models. *Economic Systems Research*, 5, 277–293.
- Lahr, M.L. (2001). Reconciling domestication techniques, the notion of re-exports and some comments on regional accounting. *Economic Systems Research*, 13, 165–179.
- Theil, H., Beerens, G.A.C., DeLeeuw, C.G. and Tilanus, C.B. (1966) *Applied Economic Forecasting*. Elsevier, New York.
- Többen, J. and Kronenberg, T. (2015) Construction of multi-regional input–output tables using the CHARM method. *Economic Systems Research*, 27, 487–507.
- Tohmo, T. (2004) New developments in the use of location quotients to estimate regional input–output coefficients and multipliers. *Regional Studies*, 38, 43–54.
- Zhao, X. and S-G. Choi (2015) On the regionalization of input–output tables with an industry-specific location quotient. *Annals of Regional Science*, 54, 901–926.

Table 1 Accuracy of estimated type I output multipliers for two South Korean regions in 2005: Zhao and Choi's findings (MAPE based on 28 sectors)

Formula	Region	
	Daegu	Gyeongbuk
SLQ	50.78	70.91
CILQ	63.01	61.85
FLQ ($\delta = 0.3$)	14.14	22.89
FLQ ($\delta = 0.4$)	9.36	15.84
FLQ ($\delta = 0.5$)	8.65	12.20
FLQ ($\delta = 0.6$)	9.91	8.48
Best single value of δ	0.5	0.6

Source: Zhao and Choi (2015, tables 8 and 9)

Minima are shown in bold type

Table 2 Reworking of Zhao and Choi's findings based on the same stated assumptions as Table 1

Formula	Region	
	Daegu	Gyeongbuk
SLQ	42.70	66.97
CILQ	45.37	56.71
FLQ ($\delta = 0.3$)	11.71	19.07
FLQ ($\delta = 0.35$)	9.74	16.40
FLQ ($\delta = 0.4$)	9.20	14.26
FLQ ($\delta = 0.45$)	9.45	13.12
FLQ ($\delta = 0.5$)	10.18	12.47
FLQ ($\delta = 0.55$)	11.15	10.91
FLQ ($\delta = 0.6$)	12.20	10.49
Best single value of δ	0.4	0.6

Table 3 Variant of Table 2 based on a type B rather than A national coefficient matrix

Formula	Region	
	Daegu	Gyeongbuk
SLQ	27.11	30.37
CILQ	27.11	26.39
FLQ ($\delta = 0.3$)	6.82	6.91
FLQ ($\delta = 0.35$)	6.46	6.45
FLQ ($\delta = 0.4$)	7.07	6.62
FLQ ($\delta = 0.45$)	8.16	7.21
FLQ ($\delta = 0.5$)	9.40	8.22
FLQ ($\delta = 0.55$)	10.44	9.79
FLQ ($\delta = 0.6$)	11.41	11.79
Best single value of δ	0.4	0.35

Table 4 Variant of Table 3 based on sectoral output rather than employment data

Formula	Region	
	Daegu	Gyeongbuk
SLQ	23.33	13.67
CILQ	20.90	18.50
FLQ ($\delta = 0.3$)	6.29	5.11
FLQ ($\delta = 0.35$)	6.12	5.44
FLQ ($\delta = 0.4$)	6.65	6.28
FLQ ($\delta = 0.45$)	7.45	7.15
FLQ ($\delta = 0.5$)	8.89	8.48
FLQ ($\delta = 0.55$)	9.92	10.43
FLQ ($\delta = 0.6$)	10.77	12.22
Best single value of δ	0.35	0.3

Table 5 Variant of Table 4 using output instead of employment to measure regional size

Formula	Region	
	Daegu	Gyeongbuk
SLQ	23.33	13.67
CILQ	20.90	18.50
FLQ ($\delta = 0.3$)	6.16	5.51
FLQ ($\delta = 0.35$)	6.89	5.13
FLQ ($\delta = 0.4$)	8.26	5.30
FLQ ($\delta = 0.45$)	9.54	5.84
FLQ ($\delta = 0.5$)	10.60	6.61
FLQ ($\delta = 0.55$)	11.71	7.41
FLQ ($\delta = 0.6$)	12.86	8.53
Best single value of δ	0.3	0.35

Table 6 Regression results based on Kowalewski's model (10)

	Kowalewski	Zhao and Choi	New results	
			Daegu	Gyeongbuk
Intercept	-0.009 (-0.08)	0.616 (17.5)	0.365 (2.74)	0.880 (6.12)
CL_j	1.266 (4.49)	10.635 (5.53)	0.541 (3.02)	-0.326 (-1.35)
SLQ_j	-0.025 (-0.38)	-0.214 (-5.45)	-0.086 (-1.66)	-0.018 (-0.41)
IM_j	-0.230 (-0.64)	3.352 (1.51)	-0.044 (-0.25)	-0.197 (-1.13)
VA_j	0.124 (1.12)	-0.247 (-0.51)	-0.253 (-1.68)	-0.830 (-3.82)
R^2	0.67	0.511	0.631	0.410
n	21	?	26	27

Source: Kowalewski (2015, table 8); Zhao and Choi (2015, table 2)

t statistics are in brackets. Sector 7 was omitted from the Daegu regression.

Table 7 New results using Kowalewski's sector-specific approach

Sector	Description	Daegu		Gyeongbuk	
		δ_j	$\hat{\delta}_j$	δ_j	$\hat{\delta}_j$
1	Agriculture, forestry and fishing	0.588	0.447	0.157	0.206
2	Mining and quarrying	0.516	0.484	0.098	0.202
3	Food, beverages and tobacco products	0.329	0.351	0.288	0.511
4	Textiles and apparel	0.353	0.209	0.297	0.498
5	Wood and paper products	0.231	0.381	0.386	0.485
6	Printing and reproduction of recorded media	0.297	0.301	0.348	0.436
7	Petroleum and coal products	0.000	0.475	0.369	0.314
8	Chemicals, drugs and medicines	0.430	0.384	0.454	0.529
9	Non-metallic mineral products	0.379	0.424	0.623	0.433
10	Basic metal products	0.404	0.440	0.611	0.482
11	Fabricated metal products except machinery and furniture	0.252	0.282	0.674	0.493
12	General machinery and equipment	0.294	0.364	0.577	0.511
13	Electronic and electrical equipment	0.358	0.442	0.660	0.439
14	Precision instruments	0.297	0.296	0.578	0.512
15	Transportation equipment	0.359	0.419	0.518	0.535
16	Furniture and other manufactured products	0.545	0.332	0.243	0.528
17	Electricity, gas, steam and water supply	0.317	0.259	0.297	0.357
18	Construction	0.297	0.218	0.607	0.449
19	Wholesale and retail trade	0.091	0.156	0.411	0.331
20	Accommodation and food services	0.221	0.196	0.564	0.498
21	Transportation	0.249	0.203	0.353	0.412
22	Communications and broadcasting	0.184	0.220	0.325	0.407
23	Finance and insurance	0.049	0.180	0.035	0.294
24	Real estate and business services	0.275	0.237	0.488	0.222
25	Public administration and defence	0.100	0.166	0.202	0.222
26	Education, health and social work	0.098	0.120	0.399	0.263
27	Other services	0.160	0.164	0.401	0.428
Mean		0.284	0.284	0.407	0.407

Source: Authors' own calculations

δ_j is the value that minimizes MAPE for the sectoral multipliers, whereas $\hat{\delta}_j$ is the estimated value from the last two regressions in Table 6. Sector 28 had to be omitted owing to missing data.

Table 8 Accuracy of estimated type I output multipliers for two South Korean regions in 2005 using different methods of estimation (MAPE based on all available sectors)

Method	Region	
	Daegu	Gyeongbuk
SLQ, Table 5	23.33	13.67
CILQ, Table 5	20.90	18.50
FLQ ($\delta = 0.3$), Table 5	6.16	5.51
FLQ ($\delta = 0.35$), Table 5	6.89	5.13
SFLQ (optimal δ_j), Table 7	1.85	2.04
SFLQ (estimated δ_j), Table 7	4.66	5.20
SFLQ (optimal δ_j), Zhao and Choi	2.885	2.121
SFLQ (estimated δ_j), Zhao and Choi	19.536	15.719

Source: Authors' own calculations based on $n = 28$ (Table 5) and $n = 27$ (Table 7); Zhao and Choi (2015, tables 4 and 5)

Table 9 Accuracy of estimated type I output multipliers for two South Korean regions in 2005 using alternative forms of Kowalewski's regression model (MAPE based on 27 sectors)

Method	Region	
	Daegu	Gyeongbuk
Linear model (10), Table 8	4.66	5.20
Semi-log model (13)	4.89	4.99
Double-log model (14)	4.72	4.52

Source: Authors' own calculations

Table 10 Characteristics of South Korean regions in 2005

	Region	Output (%)	Employment (%)	Share of imports from abroad	Share of imports from other regions	Herfindahl's index (H_r)
1	Gyeonggi-do	20.1	20.2	12.0	24.5	0.070
2	Seoul	18.2	25.4	6.0	17.3	0.112
3	Gyeongbuk	8.4	5.4	16.3	25.4	0.125
4	Gyeongsangnam-do	7.3	6.7	12.5	28.4	0.065
5	Ulsan	7.1	2.5	28.3	24.0	0.178
6	Jeollanam-do	6.5	3.3	21.9	16.3	0.123
7	Chungcheongnam-do	6.3	3.9	17.7	27.4	0.070
8	Incheon	5.5	4.8	17.1	28.8	0.058
9	Busan	5.1	7.4	7.7	26.6	0.060
10	Chungcheongsbuk-do	2.9	3.0	10.4	30.7	0.068
11	Daegu	2.9	4.7	6.1	27.9	0.061
12	Jeollabuk-do	2.7	3.2	7.4	30.4	0.067
13	Gangwon-do	2.2	2.9	4.4	23.0	0.077
14	Gwangju	2.2	2.8	9.9	30.7	0.077
15	Daejeon	1.9	2.7	6.5	28.1	0.077
16	Jeju-do	0.7	1.1	3.9	25.3	0.085
Mean		6.25	6.25	11.8	25.9	0.086
V		0.89	1.08	0.58	0.16	0.38

Source: Authors' own calculations for H_r and the share of imports

The share of imports is expressed as a proportion of gross output. V is the coefficient of variation.

Table 11 Accuracy of estimated type I output multipliers for South Korean regions in 2005 using the FLQ with different values of δ (MAPE based on 28 sectors)

	Region	Value of δ						
		0.2	0.25	0.3	0.35	0.4	0.45	0.5
1	Gyeonggi-do	14.03	11.47	9.19	7.07	5.65	4.74	4.21*
2	Seoul	6.91	6.66	6.57	6.54	6.60	6.70	6.84
3	Gyeongbuk	10.23	8.30	6.73	5.99	5.76	6.06	6.48
4	Gyeongsangnam-do	9.32	7.25	6.17	5.46	5.41	6.03	7.15
5	Ulsan	15.09	13.54	12.11	10.77	9.58	8.67	8.30*
6	Jeollanam-do	13.54	12.58	12.04	11.71	11.56	11.54	11.56
7	Chungcheongnam-do	15.94	13.28	10.55	8.41	7.09	6.65	6.66
8	Incheon	16.54	12.93	9.50	7.34	5.95	5.40	5.71
9	Busan	8.89	6.83	6.21	6.05	6.82	7.98	9.18
10	Chungcheongsbuk-do	9.72	8.51	7.72	7.59	7.80	8.65	9.97
11	Daegu	8.03	6.59	6.14	6.65	7.79	9.03	10.10
12	Jeollabuk-do	11.82	10.38	9.59	9.18	9.15	9.33	9.90
13	Gangwon-do	8.94	9.17	9.74	10.50	11.26	12.08	12.74
14	Gwangju	10.36	8.17	7.06	6.74	6.98	7.73	8.44
15	Daejeon	12.92	11.02	9.78	9.00	8.34	7.82	7.57
16	Jeju-do	10.28	10.69	11.17	11.64	11.99	12.41	12.73
Mean		11.41	9.84	8.77	8.16	7.98	8.18	8.60

Source: Authors' own calculations

* For these regions, the optimum actually occurs at $\delta > 0.5$

Table 12 Accuracy of estimated type I output multipliers for South Korean regions in 2005 based on different methods and criteria (based on 28 sectors in 16 regions)

Method	Criterion				
	MAPE	MPE	WMPE	SDSD $\times 10^3$	U
SLQ	22.224	-21.210	24.374	20.078	26.529
CILQ	23.541	-22.386	19.136	14.837	26.706
FLQ ($\delta = 0.2$)	11.411	-8.767	5.780	2.316	13.911
FLQ ($\delta = 0.25$)	9.836	-5.998	3.007	1.298	12.114
FLQ ($\delta = 0.3$)	8.768	-3.463	0.500	0.701	10.903
FLQ ($\delta = 0.325$)	8.424	-2.297	-0.642	0.552	10.538
FLQ ($\delta = 0.35$)	8.164	-1.190	-1.710	0.461	10.322
FLQ ($\delta = 0.375$)	8.022	-0.143	-2.699	0.428	10.237
FLQ ($\delta = 0.4$)	7.984	0.848	-3.615	0.435	10.256
FLQ ($\delta = 0.425$)	8.038	1.788	-4.471	0.483	10.370

Source: Authors' own calculations based on the unweighted mean of results for 16 regions

Table 13 Alternative regression models to estimate δ using data for 16 South Korean regions in 2005

	(1)	(2)	(3)	(4)
Intercept	-1.290 (-9.85)	-1.143 (-9.26)	-1.227 (-19.1)	-1.226 (-20.1)
$\ln R$	0.261 (3.65)	0.112 (1.34)	0.169 (3.87)	0.168 (4.80)
$\ln P$	0.462 (1.37)	0.361 (1.28)	0.325 (2.26)	0.325 (2.37)
$\ln I$	-2.231 (-1.41)	1.097 (0.59)	-0.024 (-0.02)	-
$\ln F$	-	0.351 (2.52)	0.316 (4.45)	0.317 (6.64)
B_{15}	-	-	0.577 (5.72)	0.577 (6.12)
R^2	0.555	0.718	0.934	0.934
AIC	-0.058	2.595	13.208	14.207
χ^2 (1) functional form	1.419	0.867	0.256	0.123
χ^2 (2) normality	4.013	19.257	0.002	0.002
χ^2 (1) heteroscedasticity	2.796	0.530	0.006	0.006

Source: Authors' own calculations

t statistics are in brackets. AIC is Akaike's information criterion. The critical values of χ^2 (1) and χ^2 (2) at the 5% level are 3.841 and 5.991, respectively.

Table 14 Alternative approaches to estimating δ for 16 South Korean regions in 2005

	δ Minimum MAPE	$\hat{\delta}$ Table 13, regression (4)	$\hat{\delta}$ Bonfiglio's method
Gyeonggi-do	0.534	0.481	-0.156
Seoul	0.337	0.336	-0.147
Gyeongbuk	0.401	0.469	0.142
Gyeongsangnam-do	0.389	0.433	0.239
Ulsan	0.542	0.543	0.129
Jeollanam-do	0.441	0.434	0.059
Chungcheongnam-do	0.470	0.472	0.240
Incheon	0.438	0.463	0.297
Busan	0.344	0.339	0.344
Chungcheongsbuk-do	0.347	0.360	0.434
Daegu	0.297	0.289	0.444
Jeollabuk-do	0.370	0.316	0.454
Gangwon-do	0.212	0.234	0.423
Gwangju	0.340	0.336	0.474
Daejeon	0.497	0.497	0.528
Jeju-do	0.196	0.191	0.522
Mean	0.385	0.387	0.277
MAPE (multipliers)	7.226	7.334	

Source: Authors' own calculations

Recent UWE Economics Papers

See <http://www1.uwe.ac.uk/bl/research/bristoleconomicanalysis> for a full list.

2016

- 1605 **Refining the application of the FLQ Formula for estimating regional input coefficients: an empirical study for South Korean regions**
Anthony T. Flegg and Timo Tohmo
- 1604 **Higher education in Uzbekistan: reforms and the changing landscape since independence**
Kobil Ruziev and Davron Rustamov
- 1603 **Circular economy**
Peter Bradley
- 1602 **Do shadow banks create money? 'Financialisation' and the monetary circuit**
Jo Michell
- 1601 **Five Safes: designing data access for research**
Tanvi Desai, Felix Ritchie and Richard Welpton

2015

- 1509 **Debt cycles, instability and fiscal rules: a Godley-Minsky model**
Yannis Dafermos
- 1508 **Evaluating the FLQ and AFLQ formulae for estimating regional input coefficients: empirical evidence for the province of Córdoba, Argentina**
Anthony T. Flegg, Leonardo J. Mastronardi and Carlos A. Romero
- 1507 **Effects of preferential trade agreements in the presence of zero trade flows: the cases of China and India**
Rahul Sen, Sadhana Srivastava and Don J Webber
- 1506 **Using CHARM to adjust for cross-hauling: the case of the Province of Hubei, China**
Anthony T. Flegg, Yongming Huang and Timo Tohmo
- 1505 **University entrepreneurship education experiences: enhancing the entrepreneurial ecosystems in a UK city-region**
Fumi Kitagawa, Don J. Webber, Anthony Plumridge and Susan Robertson
- 1504 **Can indeterminacy and self-fulfilling expectations help explain international business cycles?**
Stephen McKnight and Laura Povoledo
- 1503 **User-focused threat identification for anonymised microdata**
Hans-Peter Hafner, Felix Ritchie and Rainer Lenz
- 1502 **Reflections on the one-minute paper**
Damian Whittard
- 1501 **Principles- versus rules-based output statistical disclosure control in remote access environments**
Felix Ritchie and Mark Elliot

2014

- 1413 **Addressing the human factor in data access: incentive compatibility, legitimacy and cost-effectiveness in public data resources**
Felix Ritchie and Richard Welpton

- 1412 **Resistance to change in government: risk, inertia and incentives**
Felix Ritchie
- 1411 **Emigration, remittances and corruption experience of those staying behind**
Artjoms Ivlevs and Roswitha M. King
- 1410 **Operationalising ‘safe statistics’: the case of linear regression**
Felix Ritchie
- 1409 **Is temporary employment a cause or consequence of poor mental health?**
Chris Dawson, Michail Veliziotis, Gail Pacheco and Don J Webber
- 1408 **Regional productivity in a multi-speed Europe**
Don J. Webber, Min Hua Jen and Eoin O’Leary
- 1407 **Assimilation of the migrant work ethic**
Chris Dawson, Michail Veliziotis, Benjamin Hopkins
- 1406 **Empirical evidence on the use of the FLQ formula for regionalizing national input-output tables: the case of the Province of Córdoba, Argentina**
Anthony T. Flegg, Leonardo J. Mastronardi and Carlos A. Romero
- 1405 **Can the one minute paper breathe life back into the economics lecture?**
Damian Whittard
- 1404 **The role of social norms in incentivising energy reduction in organisations**
Peter Bradley, Matthew Leach and Shane Fudge
- 1403 **How do knowledge brokers work? The case of WERS**
Hilary Drew, Felix Ritchie and Anna King
- 1402 **Happy moves? Assessing the impact of subjective well-being on the emigration decision**
Artjoms Ivlevs
- 1401 **Communist party membership and bribe paying in transitional economies**
Timothy Hinks and Artjoms Ivlevs
- 2013**
- 1315 **Global economic crisis and corruption experience: Evidence from transition economies**
Artjoms Ivlevs and Timothy Hinks
- 1314 **A two-state Markov-switching distinctive conditional variance application for tanker freight returns**
Wessam Abouarghoub, Iris Biefang-Frisancho Mariscal and Peter Howells
- 1313 **Measuring the level of risk exposure in tanker shipping freight markets**
Wessam Abouarghoub and Iris Biefang-Frisancho Mariscal
- 1312 **Modelling the sectoral allocation of labour in open economy models**
Laura Povoledo
- 1311 **The US Fed and the Bank of England: ownership, structure and ‘independence’**
Peter Howells
- 1310 **Cross-hauling and regional input-output tables: the case of the province of Hubei, China**
Anthony T. Flegg, Yongming Huang and Timo Tohmo
- 1309 **Temporary employment, job satisfaction and subjective well-being**
Chris Dawson and Michail Veliziotis