**Online Appendix Part 1: Estimating regional imports**

A key objective of any LQ-based formula is to estimate a region’s imports from other regions. Furthermore, where the regional and national propensities to import from other countries diverge, it is also necessary to allow for that divergence. The aim of this section is to assess whether the FLQ yields more accurate adjustments for imports than does the CILQ. These formulae differ in only one respect: the FLQ makes a greater allowance for interregional imports, the smaller the relative size of a region, whereas the CILQ makes no explicit adjustment for regional size. The FLQ yields different results owing to the presence of the scalar $λ^{r}$ in the formula; see equations (4) and (5).

Now consider the following equations:

 $Δm\_{ij}^{r}=CILQ\_{ij}^{r}\_{ }\left(1-λ^{r}\right)$ if $CILQ\_{ij}^{r}<1$ (A1)

 $Δm\_{ij}^{r}=(1-CILQ\_{ij}^{r}\_{ }×λ^{r})$ if $1\leq CILQ\_{ij}^{r}<1/λ^{r}$ (A2)

 $Δm\_{ij}^{r}=0$ if $1/λ^{r}\leq CILQ\_{ij}^{r}$ (A3)

where $Δm\_{ij}^{r}$ represents the extra imports generated by using the FLQ rather than the CILQ as the regionalization technique. Also note that $m\_{ij }^{r}≡m\_{ij}^{n}+m\_{ij }^{rs},$ where $m\_{ij}^{n}$ and $m\_{ij }^{rs}$ are the propensities to import in international and interregional trade, respectively. Equation (A1) is the outcome if $CILQ\_{ij}^{r}<1 $and $FLQ\_{ij}^{r}<1$, (A2) occurs when $CILQ\_{ij}^{r}\geq 1 $and $FLQ\_{ij}^{r}<1$, while (A3) is a case where $FLQ\_{ij}^{r}\geq 1.$

 The following statistics will be employed to assess the relative performance of the FLQ and CILQ for each region *r*:

 MPE*r* = (100/27)Σ*j*($\hat{m}\_{j}^{r}- m\_{j}^{r})/m\_{j}^{r}\_{}$ (A4)

 WMPE*r* = 100Σ*j*$w\_{j }^{r}$($\hat{m}\_{j}^{r}- m\_{j}^{r})/m\_{j}^{r}\_{}$ (A5)

 WMAE*r* = Σ*j*$w\_{j }^{r}$|$\hat{m}\_{j}^{r}-m\_{j}^{r}$| (A6)

 TPE*r* = 100($\hat{M}^{r}-M^{r})/M^{r}$ (A7)

where $\hat{m}\_{j}^{r}$is the estimated propensity to import from both other regions and abroad of sector *j* in region *r* (expressed as a proportion of that sector’s gross output), whereas $m\_{j}^{r}$ is the corresponding benchmark value.$ w\_{j }^{r}$ is the proportion of regional gross output produced in sector *j*. $\hat{M}^{r}$ and $M^{r}$ are the estimated and observed total imports in region *r*. The first two statistics aim to capture any bias in estimating imports. The second formula takes into account the relative importance of sectors. The third formula minimizes the possibility that an offsetting of positive and negative errors might produce a spuriously accurate outcome. Finally, TPE (total percentage error) measures the overall error in estimating a region’s imports. A selection of results is shown in Table A1.

Looking first at the MPE results, it is striking how the CILQ understates regional propensities to import by 35.2% on average, whereas the FLQ with *δ* = 0.375 overstates them but by a more modest 7.8%. Moreover, when the results are weighted by the regional share of national output, the corresponding figures become 35.7% and 0.4%. Nevertheless, there is much interregional variation in the values of MPE generated by the FLQ. This outcome can be explained in terms of differences in the optimal values of *δ*, as shown in Table 1. With a couple of exceptions, regions exhibiting overstatement (MPE > 0) have optimal values below 0.375, whereas those with understated propensities have values above 0.375. Most of the smaller regions fall into the former category. For instance, Gangwon and Jeju have optimal deltas of 0.21, well below 0.375, and consequently have noticeably overestimated import propensities. This interregional variation in the optimal values of *δ* is discussed later.

 WMPE takes the sectoral composition of output into account and Table A1 reveals that this weighting often makes a considerable difference. For instance, for Seoul, WMPE is boosted by the predominance of sectors 19, 23 and 24 (see the list in Table 5). The gap between MPE and WMPE is typically wider for the smaller regions, which tend to have industrial structures that depart markedly from a uniform distribution. For example, in Jeju, sectors 1, 18−21 and 23−27 are over-represented in this island region, whereas sectors 2 and 4−17 are of little importance.

 When the WMPE results are averaged across regions using regional weights, one can see that the CILQ understates regional propensities to import by 30.0% on average, whereas the FLQ overstates them by 10.4%. This overstatement could be addressed, at least in part, by considering regional characteristics when selecting a value of *δ* for each region. The FLQ’s superiority over the CILQ is corroborated by the WMAE, so the favourable outcomes from both MPE and WMPE are clearly not simply due to an offsetting of negative and positive errors in estimating propensities to import. Furthermore, the TPE results reveal that the CILQ greatly understates the total imports of every region.

 **Table A1.** Assessment of accuracy of each region’s estimated imports from other South Korean regions and abroad in 2005.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Region | FLQ (*δ* = 0.375) | CILQ |
| MPE | WMPE | WMAE | TPE | MPE | WMPE | WMAE | TPE |
|  1 | Gyeonggi | −12.59 | −8.73 | 0.0421 | −12.42 | −45.32 | −44.67 | 0.1460 | −41.10 |
|  2 | Seoul | 1.73 | 29.17 | 0.0505 | −11.54 | −15.97 | −12.28 | 0.0366 | −21.45 |
|  3 | North Gyeongsang | 5.62 | 0.97 | 0.0601 | −0.76 | −35.36 | −25.54 | 0.1146 | −35.71 |
|  4 | South Gyeongsang | 3.63 | 4.64 | 0.0448 | −0.51 | −39.86 | −36.12 | 0.1409 | −37.72 |
|  5 | Ulsan | −11.85 | −7.32 | 0.0665 | −15.36 | −39.68 | −20.25 | 0.1143 | −40.33 |
|  6 | South Jeolla | −0.75 | 9.14 | 0.0450 | −11.81 | −30.20 | −11.96 | 0.0657 | −34.52 |
|  7 | South Chungcheong | −5.82 | −1.65 | 0.0516 | −8.88 | −46.75 | −39.94 | 0.1736 | −45.98 |
|  8 | Incheon | −11.77 | −8.19 | 0.0566 | −11.68 | −56.23 | −50.44 | 0.2091 | −50.79 |
|  9 | Busan | 14.47 | 29.55 | 0.0696 | 4.19 | −43.84 | −40.80 | 0.1444 | −42.29 |
|  10 | North Chungcheong | 21.20 | 18.82 | 0.0648 | 7.94 | −29.61 | −34.92 | 0.1455 | −33.90 |
|  11 | Daegu | 19.41 | 37.71 | 0.0783 | 7.98 | −36.16 | −34.98 | 0.1196 | −35.94 |
|  12 | North Jeolla | 15.34 | 29.07 | 0.0890 | 0.14 | −32.61 | −28.75 | 0.1370 | −37.96 |
|  13 | Gangwon | 34.50 | 56.72 | 0.1368 | 19.42 | −13.40 | −2.87 | 0.0529 | −16.81 |
|  14 | Gwangju | 16.73 | 29.12 | 0.0795 | 3.99 | −34.73 | −32.78 | 0.1526 | −38.03 |
|  15 | Daejeon | 12.33 | 40.95 | 0.0985 | −3.79 | −38.55 | −35.12 | 0.1548 | −41.56 |
|  16 | Jeju | 23.02 | 54.82 | 0.1235 | 2.81 | −24.72 | −13.86 | 0.0690 | −27.99 |
| Mean |   | 7.82 | 19.68 | 0.0723 | −1.89 | −35.19 | −29.08 | 0.1235 | −36.38 |
| WM |  | 0.44 | 10.42 | 0.0575 | −6.63 | −35.72 | −29.96 | 0.1175 | −36.13 |

**Online Appendix Part 2: Kowalewski’s approach**

Kowalewski posits a regression model of the following form:

 *δj* = *α + β1 CLj + β2 SLQj + β3 IMj + β4 VAj + εj* (A8)

where *CLj* is the coefficient of localization, which measures the degree of concentration of national industry *j*, *IMj* is the share of foreign imports in total national intermediate inputs, *VAj* is the share of value added in total national output and *εj* is an error term. Regional data are needed for *SLQj*, whereas *CLj*, *IMj* and *VAj* require national data. *CLj* is calculated as:

 $CL\_{j}≡0.5Σ\_{r}\left|\frac{E\_{j}^{r}}{E\_{j}^{n}}-\frac{E\_{}^{r}}{E\_{}^{n}}\right|$ (A9)

where *E* is employment. Flegg and Tohmo (2019) fitted model (A8) to data for individual South Korean regions but found that the results exhibited considerable instability. Consequently, they modified this model by imposing the restriction *β2* = 0 and re-expressing the dependent variable as the mean value of *δj* across all regions. *SLQj* was excluded as it is a region-specific variable.

 Using data for 27 sectors and 16 regions, Flegg and Tohmo obtained the following result:

 *δj* = 0.669 *+* 0.269*CLj* − 0.403*IMj* − 0.628*VAj + ej* (A10)

where *ej* is a residual. Two of the regressors (*IMj* and *VAj*) were highly statistically significant, whereas *CLj* was only marginally so. *R2* = 0.589. Using this equation, Flegg and Tohmo found that the SFLQ outperformed the FLQ in Daegu, yet the opposite was true for North Gyeongsang.**1** However, they do not give an outcome for the regions as a whole.

 Our calculations using equation (A10) gave MAPE = 7.3%, which is noticeably better than the figure of 8.0% given in Table 4 for the FLQ (with *δ* = 0.375). Thus it appears that, on average, there is a potential gain of about 0.7 percentage points from permitting *δ* to vary across sectors. This outcome is surprisingly good for a regression with an *R2* = 0.589, and it lends support to the SFLQ as a regionalization technique. Moreover, if it were possible to refine this regression and thereby obtain a better fit, then that would strengthen the argument for using the SFLQ. Even so, Flegg and Tohmo note the difficulty of finding suitable new regressors for which data would be readily available. We should also remember that the regression was fitted to data for South Korea, and that the results pertain to an ‘average’ region in that country, so one does need to be cautious when employing it in other contexts.

1. Zhao & Choi (2015) also investigated this topic using data for these two regions; for a critique of their study, see Flegg and Tohmo (2019).

**Online Appendix Part 3: regional case study**

South Gyeongsang is our chosen region. Table 1 shows that it has the fourth largest share of national output, yet the third largest share of national employment, while Figure 2 reveals that it is a coastal region bordering six other regions. Its characteristics suggest that it should be representative of the country as a whole, certainly more so than any of the metropolitan regions.

 By aggregating the transactions for the other fifteen regions, we constructed two-region I-O tables and used these to compute intraregional and interregional output multipliers. Before discussing these multipliers, however, it is instructive to examine the sources of South Gyeongsang’s intermediate inputs, as identified in Table A2. Results are shown for the FLQ and CILQ but not for the SLQ, which gave very similar outcomes to the CILQ. The sectors are defined in Table 5. WM denotes the weighted mean. Sectoral shares of total regional output were used as weights.

 Table A2 reveals that, on average, South Gyeongsang obtained 22.3% of its intermediate inputs internally and 28.4% from other regions. As expected, the CILQ greatly overstates the mean intraregional share but understates the mean interregional share, with estimates of 35.9% and 14.7%, respectively. By contrast, the FLQ yields estimates of 21.5% and 24.5%, respectively, which are much closer to the targets.

 The FLQ-based estimates of intraregional shares are mostly fairly close to the observed values. The weighted mean error is only 0.8 percentage points, representing a slight understatement. By contrast, the CILQ-based shares are highly inaccurate, with a weighted mean error of 13.6 percentage points, which constitutes a very large overstatement. Nevertheless, there are eight sectors where the FLQ-based estimates have absolute errors above five percentage points. These include Food, beverages and tobacco products (3), Transportation equipment (15), Finance and insurance (23), Public administration and defence (25), and the artificial ‘dummy’ sector 28, where the share is understated by 22.6 percentage points.

 Unlike the intraregional shares, the FLQ-based estimates of interregional shares typically fall short of the observed values. On average, these shares are understated by 3.9 percentage points. By contrast, the CILQ-based shares are understated by 13.7 percentage points and there are huge discrepancies in almost all cases between the observed and estimated shares. For the FLQ, the most problematic sectors are 4, 7−10 and 13, along with 5 and 16−17.

 Table A3 displays estimates of type I output multipliers for South Gyeongsang. It is evident that the FLQ comes close to the mean for the intraregional multipliers, whereas the CILQ greatly overstates these multipliers on average. This finding is entirely consistent with the outcomes for intraregional input shares. Furthermore, as expected, the sectors mentioned earlier where the FLQ performed less well in terms of estimating intraregional input shares are also those where the FLQ-based estimates of intraregional multipliers are the least accurate.

 As regards the interregional multipliers, it is evident that the CILQ comes very close indeed to the target on average, whereas the FLQ falls a little short. However, underlying this seemingly excellent performance of the CILQ is a worrying aspect: it occurs as a result of two offsetting sources of error. Table 11 reveals that, on average, the CILQ overstates the intraregional multipliers by 0.267, yet it also understates the interregional component of the interregional multipliers by 0.270, giving an overall error of only 0.003.**2**

 In summary, the CILQ produced strongly biased estimates of both intraregional and interregional input shares in South Gyeongsang and hence unreliable estimates of multipliers. By contrast, the FLQ gave satisfactory estimates of both intraregional input shares and output multipliers for most sectors. However, its estimates of interregional input shares and output multipliers were less satisfactory on the whole and tended to understate the true values.

1. The observed interregional component of the WM interregional multiplier is 2.019 − 1.323 = 0.696. For the CILQ, it is 2.016 − 1.590 = 0.426. Hence there is a shortfall of 0.270.

**Table A2.** Observed and estimated intraregional and interregional inputs by sector: South Gyeongsang (%).

|  |  |  |
| --- | --- | --- |
|  | Intraregional inputs | Interregional inputs |
| Observed | FLQ (*δ* = 0.375) | CILQ | Observed | FLQ (*δ* = 0.375) | CILQ |
|  1 | 15.7 | 12.3 | 24.7 | 23.9 | 19.9 | 12.3 |
|  2 | 19.2 | 15.5 | 25.9 | 25.1 | 21.6 | 14.9 |
|  3 | 33.8 | 24.2 | 49.0 | 25.3 | 25.9 | 10.8 |
|  4 | 18.2 | 21.6 | 42.9 | 36.7 | 24.2 | 8.9 |
|  5 | 22.0 | 20.3 | 39.4 | 33.2 | 25.1 | 12.9 |
|  6 | 20.4 | 34.7 | 43.9 | 29.9 | 30.6 | 18.7 |
|  7 | 9.2 | 6.2 | 6.2 | 24.6 | 9.1 | 6.9 |
|  8 | 15.7 | 18.9 | 35.9 | 40.1 | 26.5 | 15.8 |
|  9 | 23.7 | 19.2 | 38.6 | 39.6 | 27.8 | 15.8 |
|  10 | 20.4 | 23.7 | 51.2 | 44.8 | 24.3 | 6.8 |
|  11 | 23.8 | 21.7 | 32.0 | 40.8 | 34.4 | 26.8 |
|  12 | 27.7 | 24.7 | 32.0 | 34.8 | 36.0 | 30.2 |
|  13 | 18.8 | 17.9 | 36.7 | 38.6 | 19.1 | 6.9 |
|  14 | 14.3 | 14.2 | 27.6 | 34.5 | 29.3 | 20.6 |
|  15 | 24.2 | 33.7 | 47.3 | 28.1 | 30.3 | 20.9 |
|  16 | 24.4 | 34.7 | 55.7 | 36.8 | 29.8 | 11.3 |
|  17 | 12.0 | 12.7 | 18.0 | 6.0 | 15.3 | 11.5 |
|  18 | 24.4 | 25.7 | 42.1 | 26.7 | 23.3 | 12.1 |
|  19 | 21.4 | 16.7 | 32.3 | 16.4 | 18.8 | 5.8 |
|  20 | 25.2 | 32.4 | 47.2 | 27.9 | 24.7 | 14.4 |
|  21 | 10.2 | 12.4 | 20.0 | 28.4 | 24.1 | 17.6 |
|  22 | 24.2 | 20.0 | 37.8 | 24.1 | 21.6 | 7.2 |
|  23 | 21.8 | 11.5 | 24.6 | 18.6 | 14.3 | 5.9 |
|  24 | 14.9 | 16.3 | 25.9 | 10.6 | 14.8 | 6.3 |
|  25 | 22.0 | 8.8 | 16.0 | 15.3 | 13.2 | 8.8 |
|  26 | 13.2 | 10.8 | 21.0 | 16.2 | 12.4 | 6.1 |
|  27 | 23.4 | 25.9 | 41.1 | 21.9 | 24.1 | 10.8 |
|  28 | 59.6 | 37.0 | 75.6 | 33.6 | 37.7 | 13.5 |
| Mean | 21.6 | 20.5 | 35.4 | 27.9 | 23.5 | 12.9 |
| WM | 22.3 | 21.5 | 35.9 | 28.4 | 24.5 | 14.7 |

**Table A3.** Observed and estimated intraregional and interregional multipliers by sector: South Gyeongsang.

|  |  |  |
| --- | --- | --- |
|  | Intraregional multipliers | Interregional multipliers |
| Observed | FLQ (*δ* = 0.375) | CILQ | Observed | FLQ (*δ* = 0.375) | CILQ |
|  1 | 1.252 | 1.160 | 1.397 | 1.753 | 1.578 | 1.685 |
|  2 | 1.273 | 1.208 | 1.411 | 1.768 | 1.657 | 1.740 |
|  3 | 1.471 | 1.303 | 1.729 | 2.079 | 1.879 | 2.069 |
|  4 | 1.259 | 1.280 | 1.676 | 2.077 | 1.845 | 1.976 |
|  5 | 1.318 | 1.263 | 1.618 | 2.038 | 1.821 | 1.969 |
|  6 | 1.305 | 1.457 | 1.706 | 1.997 | 2.230 | 2.221 |
|  7 | 1.129 | 1.080 | 1.093 | 1.584 | 1.267 | 1.232 |
|  8 | 1.235 | 1.244 | 1.550 | 2.133 | 1.805 | 1.939 |
|  9 | 1.349 | 1.249 | 1.589 | 2.158 | 1.815 | 1.968 |
|  10 | 1.297 | 1.318 | 1.949 | 2.410 | 1.949 | 2.196 |
|  11 | 1.351 | 1.299 | 1.546 | 2.392 | 2.172 | 2.246 |
|  12 | 1.405 | 1.349 | 1.514 | 2.348 | 2.285 | 2.326 |
|  13 | 1.270 | 1.231 | 1.583 | 2.145 | 1.675 | 1.813 |
|  14 | 1.206 | 1.188 | 1.434 | 1.940 | 1.815 | 1.915 |
|  15 | 1.346 | 1.509 | 1.842 | 2.107 | 2.434 | 2.566 |
|  16 | 1.350 | 1.456 | 1.904 | 2.261 | 2.254 | 2.334 |
|  17 | 1.145 | 1.154 | 1.242 | 1.267 | 1.432 | 1.464 |
|  18 | 1.348 | 1.341 | 1.685 | 2.036 | 1.966 | 2.087 |
|  19 | 1.293 | 1.216 | 1.493 | 1.658 | 1.613 | 1.677 |
|  20 | 1.409 | 1.419 | 1.763 | 1.992 | 2.033 | 2.162 |
|  21 | 1.149 | 1.170 | 1.314 | 1.612 | 1.596 | 1.627 |
|  22 | 1.335 | 1.257 | 1.582 | 1.859 | 1.718 | 1.803 |
|  23 | 1.297 | 1.145 | 1.362 | 1.693 | 1.429 | 1.514 |
|  24 | 1.211 | 1.215 | 1.413 | 1.470 | 1.568 | 1.603 |
|  25 | 1.345 | 1.126 | 1.283 | 1.792 | 1.453 | 1.527 |
|  26 | 1.199 | 1.142 | 1.343 | 1.572 | 1.431 | 1.519 |
|  27 | 1.345 | 1.358 | 1.683 | 1.882 | 1.978 | 2.046 |
|  28 | 1.870 | 1.514 | 2.284 | 2.847 | 2.462 | 2.824 |
| Mean | 1.313 | 1.273 | 1.571 | 1.960 | 1.827 | 1.930 |
| WM | 1.323 | 1.294 | 1.590 | 2.019 | 1.907 | 2.016 |