

# A Reappraisal of the Border Effect on Relative Price Volatility

YIN-WONG CHEUNG\* & KON S. LAI\*\*

*\*University of California, Santa Cruz, CA 95064, USA and University of Hong Kong, Hong Kong and, ShanDong University, China; \*\*California State University, Los Angeles, CA 90032, USA*

**ABSTRACT** Engel & Rogers (1996) find that crossing the US–Canada border can considerably raise relative price volatility and that exchange rate fluctuations explain about one-third of the volatility increase. Using a decomposition method, this study re-evaluates the border effect. It is shown that cross-country heterogeneity in price volatility can induce a bias in measuring the border effect unless proper adjustment is made to correct it. We further examine the implication of symmetric sampling for the border effect estimation under the decomposition approach. Two conditions governing the strength of the border effect are identified. In particular, the more dissimilar the price shocks are across countries, the greater the border effect will be. Decomposition estimates also suggest that exchange rate fluctuations actually account for a large majority of the border effect.

**KEY WORDS:** Relative price volatility, exchange rate volatility, national border, distance, dissimilar shocks

**JEL CLASSIFICATION:** F31, F41

---

*Correspondence Address:* Yin-Wong Cheung, Department of Economics, University of California at Santa Cruz, California, CA 95964, USA. Email: cheung@ucsc.edu

1016-8737 Print/1743-517X Online/06/040495-19 © 2006 Korea International Economic Association  
DOI: 10.1080/10168730601027120

## Introduction

The substantial price dispersion observed between similar goods at different locations is indicative of persistent deviations from the law of one price. Engel (1993, 1999) suggest that real exchange rate variability is ascribed largely to relative price movements between similar goods across borders. Using a creative method of inter- versus intra-country analysis of intercity prices, Engel & Rogers (1996) (hereafter referred to as ER) show that the national border is an important determinant of relative price volatility even after making due allowance for the role of distance. Compared to within-country relative prices, cross-border relative prices are found to have considerably higher volatility. The finding is intriguing. The national border matters: it has a sizeable positive impact on relative price volatility.

Based on regression estimates, the ER study finds that while intercity distance accounts for some of the increase in relative price volatility, the border effect far exceeds the distance effect in magnitude. Exchange rate variability, together with price stickiness, may have played a significant role in the border effect. Price stickiness can be supplemented with local currency pricing, which severely limits exchange rate pass-through into prices (Feenstra & Kendall, 1997; Engel & Rogers, 2001b). The volatile-exchange-rate-cum-sticky-price consideration is, nonetheless, found to explain much less than half of the border effect. The border effect research has been extended in different directions in recent studies, including Engel & Rogers (2000, 2001a, 2001b), Parsley & Wei (2001), Depken & Sonora (2002), Ceglowski (2003), Cheung & Fujii (2005), and Crucini *et al.* (2005).<sup>1</sup> The fact that a rather large proportion of the border effect is still unaccounted for remains a puzzle.

This study contributes by re-evaluating the border effect in several ways. The first contribution is methodological. Using a direct decomposition method, we analyze the contributing components of the border effect. If foreign and domestic prices share the same volatility, the inter- versus intra-country analysis would yield unbiased border effect estimates. This cross-country homogeneity condition seems unnecessarily stringent and is generally violated by actual data.<sup>2</sup> When extending the analysis to general cases with heterogeneity, we show that a symmetric sampling strategy, under which the same number of cities is sampled from each country, can be used to secure unbiased decomposition estimates.

Moreover, two sufficient conditions for the border effect to occur are identified: (1) fluctuating exchange rates coupled with sticky prices, and (2) cross-border relative prices are driven by more dissimilar, as opposed to common, shocks than are within-country relative prices. The first condition is well

---

<sup>1</sup>Intercity price data have been keenly analyzed. Instead of investigating relative price volatility, some recent studies (e.g. Cecchetti *et al.*, 2002) focus on the persistence of price convergence.

<sup>2</sup>Analytically, city prices may tend to have a higher variance in one country than another due to, for example, a difference in monetary policy between the countries.

discussed in the literature, but the second is not. For the latter, dissimilar shocks, which can be micro- or macroeconomic in nature, reduce price co-movement across countries and make their relative prices more variable than those within the same country. Since both conditions are generally valid, the border effect is likely to be prevalent in cross-border price data.

The decomposition method also offers a more direct and exact way to quantify the exchange rate contribution than the regression method does. This alternative method for measuring the exchange rate contribution to the border effect can avert the collinearity problem that afflicts the regression approach used in other border effect studies. Better measurement of the exchange rate contribution – which requires accurate estimates of both the border effect and its exchange rate component – is important. It bears on the issue in the large unexplained portion of the border effect. Our decomposition estimates indicate that exchange rate fluctuations actually explain the majority of the border effect.

The rest of this paper is organized as follows. The next section examines the implication of inter-country heterogeneity in price volatility for measuring the border effect in decomposition analysis. The third section analyzes the different contributing components of the border effect and its prevalence. The fourth section verifies the validity of our analytical results based on the ER data. The fifth section reports further evidence for the significant contribution of dissimilar shocks to the border effect. The sixth section summarizes our main results.

### Cross-country Heterogeneity and the Border Effect Measurement

Cross-border relative prices between cities  $k$  and  $i$  (denoted by  $q_{t,ki}^c$ ) are given by

$$q_{t,ki}^c = p_{t,k}^d - p_{t,i}^f - e_t \text{ for } k = 1, \dots, r \text{ and } i = 1, \dots, s \quad (1)$$

where  $e_t$  is the dollar price of the foreign currency and  $p_{t,k}^d$  and  $p_{t,i}^f$  are goods prices, with the superscripts  $d$  and  $f$  indicating the US and foreign cities, respectively. Likewise, relative prices between within-country cities  $k$  and  $m$  (denoted by  $q_{t,km}^w$  with  $w = d$  and  $f$ ) are given by

$$q_{t,km}^w = p_{t,k}^w - p_{t,m}^w \text{ for } k \neq m \text{ and } k, m = 1, \dots, r \text{ or } s \quad (2)$$

All variables are expressed in terms of their first-differences in logarithms and so they are unit-free.<sup>3</sup> Equation (1) suggests a simple decomposition of the

---

<sup>3</sup>Differing slightly from the ER study, we measure the volatility of individual variables using variance instead of standard deviation to maintain consistency with our analytical results, which are derived in terms of variance as well. If  $a^2 = b^2 + c^2$ , then  $a \neq b + c$  unless either  $b = 0$  or  $c = 0$ . Hence, working with standard deviation would lead to an unexplained

volatility of cross-border relative prices:

$$\text{Var}(q_{t,ki}^c) = \text{Var}(p_{t,k}^d - p_{t,i}^f) + \text{Var}(e_t) - 2\text{Cov}(e_t, p_{t,k}^d - p_{t,i}^f). \quad (3a)$$

From equation (2), the volatility of within-country relative prices is

$$\text{Var}(q_{t,km}^w) = \text{Var}(p_{t,k}^w - p_{t,m}^w) \text{ for } w = d \text{ and } f \quad (3b)$$

Let  $\text{AVG}[\cdot]$  be the sample average of an intercity variable across relevant city pairs. Let  $\text{Var}(q_t^c)$  denote  $\text{AVG}[\text{Var}(q_{t,ki}^c)]$ , and let  $\text{Var}(q_t^w)$  denote  $\text{AVG}[\text{Var}(q_{t,km}^w)]$ . Using the average volatility of within-country relative prices as a yardstick, the ‘excess’ volatility of cross-border relative prices is measured by

$$\delta = \text{Var}(q_t^c) - \text{Var}(q_t^w) \quad (4)$$

A regression-based measure of  $\delta$  may alternatively be used, as in the ER study. When intercity relative price volatility is regressed on a constant and a border dummy variable, the border effect is captured by the regression coefficient on the border dummy. Our main analysis is based on the decomposition approach because it offers alternative insight into the causes and general applicability of the border effect and because it provides a direct measure of the contribution of exchange rate volatility to the border effect.

In measuring  $\text{Var}(q_t^c)$ , every cross-border city pair contains one US city and one foreign city, and so US and foreign cities are always sampled at the same frequency. This is not the case when computing  $\text{Var}(q_t^w)$ , however, unless the data sample has been purposely controlled such that the number of US city pairs (i.e.  $n_d$ ) equals the number of foreign city pairs (i.e.  $n_f$ ). For decomposition estimates,  $\text{Var}(q_t^w)$  is essentially a weighted average of  $\text{Var}(q_t^d)$  and  $\text{Var}(q_t^f)$ , where  $\text{Var}(q_t^d) = \text{AVG}[\text{Var}(q_{t,km}^d)]$  and  $\text{Var}(q_t^f) = \text{AVG}[\text{Var}(q_{t,ij}^f)]$ . Since  $\text{Var}(q_t^d)$  and  $\text{Var}(q_t^f)$  can have unequal values, it raises a question: Should  $n_d$  equal  $n_f$ ? Only when foreign and domestic prices share the same average volatility does the analysis, whether  $n_d$  equals  $n_f$  or not, always produce unbiased decomposition estimates of the border effect,  $\delta$ . In more general cases that allow for heterogeneity in price volatility across countries, we find that unbiased decomposition estimates may still be obtained for the border effect when  $n_d$  equals  $n_f$ .

---

discrepancy term. Similar to our study, Engel & Rogers (2001b) use variance, not standard deviation, to measure volatility.

In general, the size of possible heterogeneity bias in the decomposition estimator can be shown to be

$$\text{Bias} = \frac{n_f - n_d}{n_f + n_d} \{ \text{Var}(q_t^d) - \text{Var}(q_t^f) \} / 2 \quad (5)$$

The bias can be upward or downward, depending upon whether  $n_d$  is greater or less than  $n_f$  and whether  $\text{Var}(q_t^d)$  is larger or smaller than  $\text{Var}(q_t^f)$ . An example helps illustrate the heterogeneity bias when  $n_d \neq n_f$  (the authors owe this illustrative example to Charles Engel). To simplify computation, city prices are of equal volatility within a country, i.e.  $\text{Var}(p_{t,k}^d) = \text{Var}(p_{t,m}^d) = \sigma_d^2$  and  $\text{Var}(p_{t,i}^f) = \text{Var}(p_{t,j}^f) = \sigma_f^2$ . We also consider that the correlation coefficient  $\rho$  for price changes is the same for every cross-border city pair (i.e.  $\text{Cov}(p_{t,k}^d, p_{t,i}^f) = \rho_c \sigma_d \sigma_f$ ), that the correlation coefficient for price changes is the same for every within-country pair (i.e.  $\text{Cov}(p_{t,k}^d, p_{t,m}^d) = \rho_w \sigma_d^2$  and  $\text{Cov}(p_{t,i}^f, p_{t,j}^f) = \rho_w \sigma_f^2$ ), and that the net exchange rate effect for every cross-border pair is given by  $\xi = \text{Var}(e_t) - 2\text{Cov}(e_t, p_{t,k}^d - p_{t,i}^f)$ . Accordingly,

$$\text{Var}(q_t^c) = \text{Var}(q_{t,ki}^c) = \sigma_d^2 + \sigma_f^2 - 2\rho_c \sigma_d \sigma_f + \xi \quad (6a)$$

$$\text{Var}(q_t^d) = \text{Var}(p_{t,k}^d - p_{t,m}^d) = 2\sigma_d^2 - 2\rho_w \sigma_d^2 \quad (6b)$$

$$\text{Var}(q_t^f) = \text{Var}(p_{t,i}^f - p_{t,j}^f) = 2\sigma_f^2 - 2\rho_w \sigma_f^2 \quad (6c)$$

So long as  $n_d = n_f$ , the same estimate of the volatility increase will correctly be obtained:

$$\delta = \text{Var}(q_t^c) - \{ \text{Var}(q_t^d) + \text{Var}(q_t^f) \} / 2 = \rho_w \sigma_d^2 + \rho_w \sigma_f^2 - 2\rho_c \sigma_d \sigma_f + \xi \quad (7)$$

If skewed sampling (i.e.  $n_d \neq n_f$ ) is applied, however, the volatility increase will be estimated as

$$\delta_{SK} = \text{Var}(q_t^c) - \{ n_d \text{Var}(q_t^d) + n_f \text{Var}(q_t^f) \} / (n_d + n_f) \quad (8)$$

which produces an estimation bias of the following size:

$$\text{Bias} = \delta_{SK} - \delta = \frac{n_f - n_d}{n_f + n_d} (1 - \rho_w) (\sigma_d^2 - \sigma_f^2) \quad (9)$$

The presence of deviations from the law of one price implies that  $\rho_w < 1$ . Hence,  $\delta_{SK} \neq \delta$  unless  $\sigma_d^2$  just happens to equal  $\sigma_f^2$ . The border effect estimate can be increased or decreased systematically by varying the relative numbers of foreign and US cities sampled. If US prices were more (less) volatile than

foreign prices, sampling more US than foreign cities would underestimate (overestimate) the border effect unless we can find a proper way to correct the bias.

Leaving out the apparent complexity of economic analysis, the issue may be applicable to inter- versus intra-group analysis involving heterogeneous populations in general. Skewed samples may introduce a heterogeneity bias in estimating cross-group effects when pairings are drawn from populations with unequal means or variances (see the Appendix for a non-economic example) unless proper adjustment is made to correct or eliminate it.

### Further Decomposition of the Border Effect

We next turn to questions about the prevalence of the border effect and its contributing components: should the border effect be commonly observed in all cross-border data? If so, what contributes to the general existence of excess relative price volatility? How much of the border effect is attributable to exchange rate fluctuations?

The border effect can be broken into two parts – one relates to the exchange rate and one does not:

$$\delta = \xi + \pi \quad (10)$$

where  $\xi = \text{Var}(e_t) - 2\text{AVG}[\text{Cov}(e_t, p_{t,k}^d - p_{t,i}^f)]$  and  $\pi = \text{AVG}[\text{Var}(p_{t,k}^d - p_{t,i}^f)] - \text{AVG}[\text{Var}(p_{t,k}^w - p_{t,m}^w)]$ . Let  $\text{Var}(p_t^f) = \text{AVG}[\text{Var}(p_{t,i}^f)]$  and  $\text{Var}(p_t^d) = \text{AVG}[\text{Var}(p_{t,k}^d)]$ . The non-exchange rate component  $\pi$  can be further rewritten such that:

$$\delta = \xi + \eta + v \quad (11)$$

where  $\eta + v = \pi$  and

$$\eta = 2\{\text{AVG}[\text{Cov}(p_{t,k}^w, p_{t,m}^w)] - \text{AVG}[\text{Cov}(p_{t,k}^d, p_{t,i}^f)]\} \quad (12)$$

$$v = \frac{n_f - n_d}{n_f + n_d} \{\text{Var}(p_t^f) - \text{Var}(p_t^d)\}. \quad (13)$$

The last term  $v$ , which captures part of the heterogeneity bias noted in equation (5), exists only for skewed samples.<sup>4</sup> When  $n_d = n_f$ ,  $v = 0$  and so  $\eta = \pi$ . Hence, for non-skewed samples,  $\delta = \xi + \eta$  exactly. The validity of this adding-up decomposition condition can be checked and verified empirically.

---

<sup>4</sup>It can further be shown that the estimation of  $\eta$  is also biased in the presence of skewed samples. When  $n_d \neq n_f$ ,  $v$  plus the bias in the  $\eta$  estimate will equal the total bias in the  $\delta$  estimate.

For skewed samples, on the other hand,  $\delta \neq \xi + \eta$ . In cases of  $n_d > n_f$ , for instance, we will obtain that  $\delta < \xi + \eta$  when  $\text{Var}(p_t^d) > \text{Var}(p_t^f)$ , but that  $\delta > \xi + \eta$  when  $\text{Var}(p_t^d) < \text{Var}(p_t^f)$ . The former case signifies underestimation of  $\delta$ , whereas the latter case signifies overestimation of  $\delta$ . As demonstrated in the previous section, when domestic prices are more (less) volatile than foreign prices, use of  $n_d > n_f$  underestimates (overestimates)  $\delta$ . In turn, underestimating  $\delta$  will result in  $\delta < \xi + \eta$ , while overestimating  $\delta$  will lead to  $\delta > \xi + \eta$ . Nevertheless, except when severe underestimation of  $\delta$  occurs, whether the border effect on relative price volatility is positive or not is determined ultimately by the values of  $\xi$  and  $\eta$ .

### *On the Prevalence of the Border Effect*

The simple decomposition can help assess – and bring out the general applicability of – the border effect. As the decomposition shows, if  $\xi$  and  $\eta$  are both positive, increased volatility will occur in cross-border relative prices. The  $\xi$  component indicates the net exchange rate effect. It incorporates not only exchange rate variability but also short-run co-movement between exchange rates and prices. The latter recognizes possible adjustment of prices in response to exchange rate changes. Since prices are sticky in the short run and exchange rate pass-through is far from complete, the covariance term is rather small in size compared to  $\text{Var}(e_t)$ . Since  $\text{Var}(e_t)$  is always positive,  $\xi$  should be positive as well.

Like  $\xi$ ,  $\eta$  is typically positive. It represents the differential between the covariation of prices within the same country and that of prices across countries. In general, intercity prices can be considered driven by a combination of dissimilar (weakly correlated) and common (strongly correlated) shocks. The more the cities are subject to dissimilar – as opposed to common – shocks, the weaker the intercity price co-movement. There are good reasons for significant dissimilar shocks to prevail between countries. For instance, usual macroeconomic shocks (e.g. changes in money supply, federal taxes, government spending, or productivity) are country-specific. Such shocks may act as common shocks to cities within the same country, but as dissimilar shocks to cities of different countries. Some microeconomic shocks (e.g. changes in tastes or transportation costs) also tend to be more similar – and induce more correlated price changes – for cities within the same border than for cities separated by a border. All in all, the border matters because cross-border relative prices are likely subject to more dissimilar and less common shocks than are within-country relative prices. Within-country price covariation should thus be stronger than cross-border price covariation, and so  $\eta > 0$ . We will refer to this as the differential price shock effect.<sup>5</sup>

---

<sup>5</sup>When adapted to the illustrative example in the second section,  $\eta$  is given by  $\rho_w \sigma_d^2 + \rho_w \sigma_f^2 - 2\rho_c \sigma_d \sigma_f$ . Since  $\sigma_d^2 + \sigma_f^2 \geq 2\sigma_d \sigma_f$ , a sufficient condition for the differential shock effect (i.e. for  $\eta > 0$ ) is that  $\rho_w > \rho_c$ , meaning simply that within-country prices are more correlated than are cross-border prices.

A Monte Carlo experiment can illustrate that dissimilar price shocks across countries can generate excess relative price volatility. To simplify, both the home country and the foreign country are considered to have two cities, with the following data generating processes for price changes:

$$p_{t,k}^d = a_k + \varepsilon_{t,k}^d, k = 1, 2 \quad (14a)$$

$$p_{t,i}^f = b_i + \varepsilon_{t,i}^f, i = 1, 2 \quad (14b)$$

where  $a_k$  and  $b_i$  are some constants and  $\varepsilon_{t,k}^d$  and  $\varepsilon_{t,i}^f$  denote price innovations. The innovation variance is standardized to equal unity so that  $Var(\varepsilon_{t,k}^d) = 1 = Var(\varepsilon_{t,i}^f)$ . Let the cross-border correlation of price shocks ( $\varepsilon_{t,k}^d$  and  $\varepsilon_{t,i}^f$ ) be  $\rho_c$ . Let the within-country correlation of price shocks be the same in both countries such that  $Cov(\varepsilon_{t,1}^d, \varepsilon_{t,2}^d) = \rho_w = Cov(\varepsilon_{t,1}^f, \varepsilon_{t,2}^f)$ . If price shocks are more (less) dissimilar for cross-border city pairs than for within-country pairs, then  $\rho_c < \rho_w$  ( $\rho_c > \rho_w$ ). To focus the analysis on the differential shock effect, the exchange rate component, which is being treated as fixed, is omitted.

In the simulation experiment, we arbitrarily set  $\rho_w$  equal to 0.5 and  $\rho_c$  equal to 0.1, 0.3, 0.5, 0.7 or 0.9. The simple design covers different interesting situations, including  $\rho_c < \rho_w$ ,  $\rho_c = \rho_w$ , and  $\rho_c > \rho_w$ . In each replication, excess relative price volatility  $\delta$  is computed. Simulation results are obtained based on 10,000 replications, and they are reported in Table 1. The results show that the size of the border effect is determined systematically by the value of  $\rho_c$  relative to  $\rho_w$ . In the case of main interest here, when  $\rho_c < \rho_w$  (i.e. when shocks to cross-border prices are less correlated than those to within-country prices), the  $\delta$  estimate is significantly greater than zero. The smaller the value of  $\rho_c$ , the greater the border effect. The results confirm that when price shocks are less correlated between countries than within countries, the differential impact of such shocks on cross-border prices raises relative price volatility.

In sum, the sufficient conditions for the border effect to occur are: (i) volatile exchange rates operating under sticky prices, and (ii) cross-border relative prices are driven by more dissimilar shocks (and less common shocks) than are within-country relative prices. The former ensures a positive exchange rate effect on relative price volatility, while the latter produces a positive differential shock effect on relative price volatility.<sup>6</sup> With these two conditions being generally satisfied in practice, the border effect on relative price volatility will likely exist widely in cross-border data. All else being equal, the more significant the dissimilar shocks across countries, the greater the border effect.

---

<sup>6</sup>While  $\xi > 0$  and  $\eta > 0$  are sufficient conditions, the necessary and sufficient condition for the existence of the border effect is weaker, requiring only that  $\xi + \eta > 0$ .



**Table 1.** Monte Carlo results for the border effect generated by dissimilar shocks

	Mean estimate of $\delta$	Standard deviation of $\delta$
$\rho_c = 0.1 < \rho_w$	0.800	0.135
$\rho_c = 0.3 < \rho_w$	0.401	0.097
$\rho_c = 0.5 = \rho_w$	0.001	0.061
$\rho_c = 0.7 > \rho_w$	-0.400	0.037
$\rho_c = 0.9 > \rho_w$	-0.820	0.025

*Notes:* The border effect on relative price volatility is measured by  $\delta$ .  $\rho_c$  indicates the correlation of price shocks for cross-border city pairs, and  $\rho_w$  indicates the correlation of price shocks for within-country city pairs ( $\rho_w$  is set equal to 0.5 in simulation). In the cases where  $\rho_c < \rho_w$ , price shocks are more dissimilar for cross-border city pairs than for within-country pairs. In the cases where  $\rho_c > \rho_w$ , price shocks are less dissimilar for cross-border city pairs than for within-country pairs. All the simulation estimates reported are based on 10,000 replications.

### *Measuring the Contribution of Exchange Rate Fluctuations*

The volatility decomposition does not just help establish the generality of the border effect result, but also provides a direct, simple way to measure the exchange rate contribution to the border effect. The importance of the exchange rate contribution has been a debated issue. Previous border effect studies commonly use a dummy variable to capture the border effect in regression analysis. Due to a collinearity problem between the exchange rate variable and the border variable, the contribution of exchange rate volatility cannot be measured directly using the regression method.

The decomposition technique, by contrast, circumvents the collinearity problem. It provides a straightforward method for measuring the exchange rate contribution. Specifically, the exchange rate contribution to the border effect can be computed as a simple ratio:  $Var(e_t)/\delta$  in gross term or  $\xi/\delta$  in net term. The  $Var(e_t)/\delta$  measure gauges the contribution in terms of exchange rate volatility only. The  $\xi/\delta$  measure is similar to the  $Var(e_t)/\delta$  measure, but it partials out possible interaction effects between exchange rate and price changes, including exchange rate pass-through into prices. In this way, the  $\xi/\delta$  measure gives the net exchange rate contribution to the border effect. Either measure underscores the importance of obtaining unbiased decomposition estimates of the border effect.

### **Decomposition Estimates of the Contributing Components of the Border Effect**

To illustrate, some decomposition estimates of the border effect are obtained based on the ER dataset. It considers Canadian (CN) and US consumer prices for 14 different categories of goods, hereafter labeled by  $\{G-1, G-2, \dots, G-14\}$ , which comprise about 95% of consumer expenditures. The data

cover the period from June 1978 through December 1994, and the whole dataset includes a total of 9 CN and 14 US cities. Monthly price data are available for all the CN cities (Calgary, Edmonton, Montreal, Ottawa, Quebec, Regina, Toronto, Vancouver and Winnipeg) and four 'core' US cities (Chicago, Philadelphia, Los Angeles, and New York). For five additional US cities (Dallas, Detroit, Houston, Pittsburgh and San Francisco), price data are given in even-numbered months. For five other US cities (Baltimore, Boston, Miami, St. Louis and Washington DC), data are available for odd-numbered months. All price changes are percentage changes calculated as two-month differences in logarithms of price indices.

With the sample (nine CN and 14 US cities) being skewed toward US cities, equations (11) and (13) predict that  $\delta < \xi + \eta$  when  $\text{Var}(p_t^d) > \text{Var}(p_t^f)$ , but that  $\delta > \xi + \eta$  when  $\text{Var}(p_t^d) < \text{Var}(p_t^f)$ . The former indicates the decomposition measure underestimates  $\delta$ , whereas the latter indicates it overestimates  $\delta$ .

Panel A of Table 2 contains estimates of  $\delta$ ,  $\xi$  and  $\eta$ . Results are given for individual goods indices. As shown, the estimates of  $\xi$  and  $\eta$  invariably confirm that the contributions from the exchange rate effect and the differential price shock effect are both positive. In all cases,  $\delta > 0$  and there is increased volatility in cross-border relative prices. Comparing estimates between  $\delta$ ,  $\xi$  and  $\eta$  reveals exactly the same pattern predicted by theoretical analysis. For the eight goods indices with  $\text{Var}(p_t^d) > \text{Var}(p_t^f)$ , we have  $\delta < \xi + \eta$ . For the six other goods indices with  $\text{Var}(p_t^d) < \text{Var}(p_t^f)$ , we have  $\delta > \xi + \eta$ . Averaging across all 14 goods indices, the volatility of US prices is almost three times higher than that of CN prices.

As an interesting contrast, we examine an alternative, non-skewed sample that includes nine CN and nine US cities only. The earlier data sample includes all 14 US cities with a combined use of even-month and odd-month data. To adopt symmetric sampling, odd-month data for five US cities are dropped from the sample here. The new sample now contains 36 CN–CN pairs and 36 US–US pairs, along with 81 CN–US pairs. Given that US–US and CN–CN city pairs are equal in number, there is no bias in the decomposition estimate of the border effect. With  $n_d = n_f$ , the theoretical analysis predicts the adding-up decomposition condition,  $\delta = \xi + \eta$ , to prevail, independent of whether  $\text{Var}(p_t^d)$  is larger or less than  $\text{Var}(p_t^f)$ .

As shown by Panel B of Table 2, the theoretical prediction is totally borne out by the empirical estimates from the non-skewed sample. All the estimates of  $\delta$ ,  $\xi$  and  $\eta$  are positive for individual goods indices.<sup>7</sup> The contributions of  $\xi$  and  $\eta$  also add up exactly to the border effect in every case, satisfying the adding-up condition for unbiased decomposition estimates. On average, the non-skewed sample – as predicted – yields a higher border effect than the skewed sample does. In relative terms, the border effect on relative

---

<sup>7</sup>Consistent with sticky prices, the covariance between exchange rates and prices was found to be small in magnitude (less than 4% on average in relative magnitude) compared with exchange rate variance.

Table 2. Estimation of the border effect and its breakdown components

Goods index	$Var(p_t^f)$	$Var(p_t^d)$	$\delta$	$\xi$	$\eta$	Adding-up condition	Exchange rate contribution to $\delta$	
							Gross: $Var(e_t)/\delta$	Net: $\xi/\delta$
(A) Skewed sample: $n_f = 36 < n_d = 66$								
G-1	4.09	1.91	3.38	1.89	0.73	$\delta > \xi + \eta$	74.8%	56.0%
G-2	1.21	1.06	3.20	2.64	0.43	$\delta > \xi + \eta$	79.1%	82.6%
G-3	2.56	2.50	4.31	3.01	1.20	$\delta > \xi + \eta$	58.6%	69.7%
G-4	0.83	3.20	3.09	2.70	1.02	$\delta < \xi + \eta$	82.0%	87.5%
G-5	6.83	19.41	6.65	1.77	8.99	$\delta < \xi + \eta$	38.0%	26.6%
G-6	0.91	2.43	2.48	2.60	0.23	$\delta < \xi + \eta$	102.1%	104.8%
G-7	3.02	17.77	4.53	2.18	6.50	$\delta < \xi + \eta$	55.8%	48.0%
G-8	3.66	67.92	13.82	1.94	30.00	$\delta < \xi + \eta$	18.3%	14.1%
G-9	3.91	23.83	1.14	2.44	1.35	$\delta < \xi + \eta$	221.1%	213.3%
G-10	3.47	3.05	4.71	2.29	2.20	$\delta > \xi + \eta$	53.7%	48.7%
G-11	24.79	15.57	24.44	2.46	18.28	$\delta > \xi + \eta$	10.3%	10.1%
G-12	3.45	1.05	4.41	2.91	0.69	$\delta > \xi + \eta$	57.3%	65.9%
G-13	1.73	3.56	2.35	2.42	0.20	$\delta < \xi + \eta$	107.6%	103.0%
G-14	0.79	2.27	2.54	2.49	0.28	$\delta < \xi + \eta$	99.5%	98.1%
Average	4.37	11.82	5.63	2.41	5.15	$\delta < \xi + \eta$	75.6%	73.5%
(B) Non-skewed sample: $n_f = 36 = n_d$								
G-1	4.09	1.74	2.99	1.94	1.05	$\delta = \xi + \eta$	89.4%	64.8%
G-2	1.21	0.88	3.43	2.87	0.56	$\delta = \xi + \eta$	78.0%	83.7%
G-3	2.56	2.35	4.50	3.11	1.39	$\delta = \xi + \eta$	59.5%	69.0%
G-4	0.83	3.44	4.05	3.00	1.05	$\delta = \xi + \eta$	66.1%	74.0%
G-5	6.83	20.21	9.68	2.00	7.69	$\delta = \xi + \eta$	27.6%	20.6%
G-6	0.91	2.28	3.15	2.81	0.33	$\delta = \xi + \eta$	85.0%	89.5%
G-7	3.02	14.66	7.64	1.83	5.81	$\delta = \xi + \eta$	35.0%	24.0%
G-8	3.66	51.81	15.26	1.27	13.99	$\delta = \xi + \eta$	17.5%	8.3%
G-9	3.91	20.84	3.55	2.58	0.97	$\delta = \xi + \eta$	75.4%	72.8%
G-10	3.47	3.31	4.61	2.48	2.13	$\delta = \xi + \eta$	57.9%	53.8%
G-11	24.79	15.98	25.93	2.14	23.79	$\delta = \xi + \eta$	10.3%	8.2%
G-12	3.45	0.96	4.08	3.16	0.92	$\delta = \xi + \eta$	65.5%	77.4%
G-13	1.73	3.37	3.02	2.64	0.38	$\delta = \xi + \eta$	88.5%	87.5%
G-14	0.79	2.43	3.12	2.73	0.39	$\delta = \xi + \eta$	85.7%	87.6%
Average	4.37	10.30	6.69	2.47	4.32	$\delta = \xi + \eta$	60.1%	58.7%

Notes:  $n_f$  indicates the number of CN–CN pairs, and  $n_d$  is the number of US–US pairs for each goods index.  $Var(p_t^f)$  indicates the average volatility of foreign prices, whereas  $Var(p_t^d)$  indicates the average volatility of US prices. For the volatility calculation, all price changes are measured as percentage changes. The column ' $\delta$ ' gives estimates of excess volatility in cross-border relative prices. When underestimation of  $\delta$  occurs,  $\delta < \xi + \eta$ . When overestimation of  $\delta$  occurs,  $\delta > \xi + \eta$ . The columns ' $Var(e_t)/\delta$ ' and ' $\xi/\delta$ ' indicates the gross and net contributions of the exchange rate to the border effect, respectively.

price volatility is given by  $\delta/Var(q_t^w)$ . Based on our calculation,  $\delta/Var(q_t^w)$  is on average equal to 0.771 for the non-skewed sample as opposed to 0.496 for the skewed sample. Accordingly, crossing the border on average raises relative price volatility by 77.1% based on the non-skewed sample.

Based on indirect approximation measurements, the ER study estimates the average contribution of exchange rate volatility to be about 30% of the border effect, implying that most of the border effect comes from non-exchange rate sources. This estimate is surprisingly low. When prices are sticky, cross-border relative price movement is expected to reflect in large part exchange rate changes. Decomposition estimates of the exchange rate contribution are reported in the last two columns of Table 2. In 10 of the 14 cases, the exchange rate explains more than half of the border effect. For the skewed sample, exchange rate volatility explains, on average, 73.5% of the border effect.<sup>8</sup> For the non-skewed sample, the average exchange rate contribution is calculated to be 60.1%. Even after adjusting for possible effects of exchange rate changes on prices, the net exchange rate contribution is computed to be 58.7% on average.

### **Additional Evidence on the Differential Price Shock effect**

In accord with our decomposition analysis, the simulation results have earlier shown that when price shocks are more dissimilar across the border than within the same country, they result in higher volatility in cross-border relative prices. The reported  $\eta$  estimates are invariably positive, and they are indicative of the differential shock effect. To seek additional evidence for the impact of dissimilar shocks, the innovation part of price changes is identified and isolated from their deterministic part by fitting a high-order autoregressive process with seasonal dummies to price data for each city. The lag order is determined using the standard Akaike information criterion.<sup>9</sup> The covariance of price innovations is constructed for every city pair in the non-skewed sample. The more dissimilar the price shocks, the weaker the covariation of shocks. If dissimilar shocks are significant across countries, the covariance of cross-border price shocks should be much lower than within-country shocks. Average covariance estimates for price innovations are given in Table 3, and they bear out the dissimilar-shock proposition. For all the goods indices, shocks to cross-border relative prices are much more dissimilar than shocks to within-country relative prices.

To provide additional confirmatory evidence on the effect of dissimilar price shocks, regression analysis is further conducted for each goods index based on the non-skewed sample. In our earlier decomposition analysis, the use of symmetry sampling is crucial for removing the possible heterogeneity bias in estimating the exchange rate effect. In regression analysis, alternative

---

<sup>8</sup>Seemingly, the decomposition estimate of the exchange rate contribution computed for the skewed data sample may even exceed the border effect in a number of disaggregate cases. This situation is misleading and can happen only when the border effect is greatly underestimated.

<sup>9</sup>A fixed third-order autoregressive specification, used by the ER study to model price processes, was also tried. Our results were found to be robust with respect to lag selection.

Table 3. Intercity covariance estimates for price shocks

Goods index	Average intercity covariance of price shocks	
	Within-country pairs	Cross-border pairs
G-1	0.826	0.647
G-2	0.280	0.023
G-3	0.703	0.258
G-4	0.272	0.082
G-5	1.070	0.441
G-6	0.179	0.063
G-7	0.729	0.202
G-8	1.585	0.411
G-9	0.905	0.838
G-10	1.406	0.844
G-11	6.533	1.465
G-12	0.101	0.061
G-13	0.269	0.130
G-14	0.203	0.104
Average estimate	1.076	0.398

Notes: Price shocks are constructed as the innovations estimated from fitting an autoregressive process to the relevant price series for each city, with the lag order deter-

tools such as dummy variables may be used to account for inter-country heterogeneity. On the other hand, the regression analysis does not determine the exchange rate contribution to the border effect.

In our data, there are 72 within-country (36 US–US and 36 CN–CN) city pairs and 81 cross-border city pairs, yielding a total of 153 observations. The regression equation is specified as:<sup>10</sup>

$$\begin{aligned}
 RPV_{j,k} = & \varphi_1 DIST_{j,k} + \varphi_2 BD_{j,k} + \varphi_3 SCV_{j,k} + \varphi_4 NER_{j,k} \\
 & + \sum_{r=1}^n \pi_r CD_r^{US} + \sum_{s=1}^n \psi_s CD_s^{CN} + \varepsilon_{j,k}
 \end{aligned} \tag{15}$$

where  $RPV_{j,k}$  represents the relative price volatility for cities  $j$  and  $k$ ,  $DIST_{j,k}$  is the log of the distance between the cities for the city pair,  $BD_{j,k}$  is the border dummy variable ( $BD_{j,k} = 1$  when a border exists between cities  $j$  and  $k$ ;  $BD_{j,k} = 0$  otherwise),  $SCV_{j,k}$  represents the covariance of price shocks for

<sup>10</sup>The regression takes account of the possible distance effect on relative price volatility. Geographic distance between locations is a natural barrier that breeds market segmentation. Goods arbitrage is more costly the further apart are the cities. By hindering goods arbitrage, distance may weaken the market linkage of price changes between cities in distant locations, thereby making relative price changes more variable.

cities  $j$  and  $k$ ,  $NER_{j,k}$  captures the net exchange rate effect ( $NER_{j,k} = Var(e_t) - 2Cov(e_t, p_{t,j} - p_{t,k})$  for cross-border city pairs;  $NER_{j,k} = 0$  otherwise), and  $\varepsilon_{j,k}$  is the random error. Dummy variables for individual cities ( $CD_r^{US} = 1$  when  $r = j$  or  $k$  and  $= 0$  otherwise, and  $CD_s^{CN} = 1$  when  $s = j$  or  $k$  and  $= 0$  otherwise) are included with  $n$  being the number of cities sampled from each country. These city dummies allow price variance to vary across cities. The coefficient for the exchange rate variable,  $\varphi_4$ , will be imposed as unity – in accord with the decomposition equation (3a) – to avert the collinearity problem with the border dummy. In the presence of other explanatory variables, the border dummy captures the unexplained part of the border effect. To the extent that distance and the border both raise relative price volatility, we expect to find  $\varphi_1 > 0$ ,  $\varphi_2 > 0$ . On the other hand, if dissimilar shocks cause higher relative price volatility, we expect to find  $\varphi_3 < 0$ .

Table 4 reports the regression results for individual goods indices (coefficient estimates for the 18 city dummy variables are not reported to conserve space). These results are generally consistent with those from our earlier decomposition analysis. The regression equation fits the data extremely well, with an adjusted  $R^2$  value of 0.97 on average. Interestingly, the adjusted  $R^2$  value can remain similarly high even if the border dummy is omitted. This suggests that the unexplained portion of the border effect becomes small once the other explanatory variables (in particular, the covariation of price shocks and the variability of the nominal exchange rate) are included. Moreover, even after controlling for the distance effect, the price-shock coefficient is strongly significant and has the correct negative sign in every case.<sup>11</sup> Consistent with the proposition of differential shock effects, the results indicate that the more dissimilar the price shocks, the greater will be the relative price volatility.

In the actual method used by ER, the two groups of city dummy variables can completely correct the potential bias in estimating the border effect. That is, the heterogeneity bias we identify in decomposition analysis does not exist in the ER or any other studies that used city dummies. Each group of city dummies functions together as, in effect, a country dummy variable: one for those city pairs that contain a US city, and another for city pairs that include a Canadian city. These two groups of city dummies can thus account for the possible difference in number between US–US and CN–CN city pairs. A recent study by Gorodnichenko & Tesar (2005) also identifies a similar problem in border effect estimation induced by inter-country heterogeneity in price variance. To address the heterogeneity problem, Gorodnichenko & Tesar (2005) propose an alternative method through the use of country-specific dummy variables.

---

<sup>11</sup>The ER study suggests an interesting way to gauge the size of the border effect as equivalent geographic distance. Relative to the border coefficient, a small distance coefficient will imply a large estimate of the border width. As shown in Table 4, however, the distance coefficient is not statistically significant in 10 out of the 14 cases.

Table 4. Regression results explaining relative price volatility

Goods index	Coefficient estimates and significance				BD omitted	
	DIST	BD	SCV	NER	Adj. R <sup>2</sup>	Adj. R <sup>2</sup>
G-1	0.027 (0.048)	0.383 (0.054)**	-2.174 (0.136)**	1.000	0.98	0.98
G-2	-0.005 (0.015)	0.208 (0.057)**	-1.484 (0.148)**	1.000	0.99	0.99
G-3	-0.011 (0.039)	0.773 (0.098)**	-1.684 (0.136)**	1.000	0.98	0.97
G-4	0.167 (0.079)*	0.689 (0.148)**	-1.200 (0.495)**	1.000	0.96	0.95
G-5	0.787 (0.799)	5.669 (0.929)**	-3.161 (0.619)**	1.000	0.90	0.87
G-6	0.003 (0.021)	0.111 (0.045)**	-1.904 (0.181)**	1.000	0.99	0.99
G-7	0.128 (0.203)	3.710 (0.374)**	-1.990 (0.306)**	1.000	0.97	0.96
G-8	0.843 (0.713)	10.180 (1.772)**	-2.155 (0.623)**	1.000	0.94	0.93
G-9	0.212 (0.194)	0.800 (0.265)**	-2.024 (0.362)**	1.000	0.99	0.99
G-10	0.218 (0.065)**	(0.139)**	-1.018 (0.205)**	1.000	0.98	0.94
G-11	1.297 (0.624)*	9.613 (2.767)**	-2.148 (0.489)**	1.000	0.92	0.90
G-12	0.015 (0.055)	0.963 (0.069)**	-1.770 (0.280)**	1.000	0.99	0.99
G-13	-0.004 (0.029)	0.109 (0.045)**	-1.946 (0.090)**	1.000	0.99	0.99
G-14	0.064 (0.018)**	0.230 (0.041)**	-1.431 (0.193)**	1.000	0.99	0.99
Average coefficient	0.267	2.494	-1.864	1.000	0.97	0.96

Notes: The dependent variable is the intercity relative price volatility (*RPV*), and the explanatory variables include *DT* (intercity distance, measured in logarithms and in miles), *BD* (the border dummy), *SCV* (intercity covariance of price innovations), *NER* (the net exchange rate effect), and 18 city dummy variables. The coefficient on *NER* is imposed as unity based on the direct decomposition result. The coefficient estimates for all the 18 city dummy variables are omitted to conserve space. There are 153 observations in each regression, and the data are based on the non-skewed sample with nine Canadian and nine US cities. Heteroskedasticity-consistent standard errors are reported in parentheses. Statistical significance is indicated by double asterisks (\*\*) for the 1% level and a single asterisk (\*) for the 5% level. The column 'Adj.  $R^2$ ' gives the adjusted coefficient of determination for the corresponding regression.

All in all, the potential heterogeneity bias in measuring the border effect can either be corrected in regression analysis using city or country dummy variables, or can be averted altogether using a symmetric sampling method. The symmetric sampling method is crucial and necessary to our decomposition analysis for which the method of dummy variables cannot be used. In contrast to the regression analysis, which cannot overcome the collinearity between the border dummy and the exchange rate variable, the more direct decomposition analysis is able to provide precise estimates of the exchange rate contribution to the border effect.

Finally, we recognize that dissimilar price shocks may possibly reflect some systematic patterns of intercity economic changes. In particular, economic factors that can influence intercity price co-movement can be potential contributors. Except for wages, city-level data on cost and demand factors relevant to our analysis are not readily available, especially for Canadian cities. The ER study notes that national labor markets separated by a border should be less integrated than local labor markets within a country. We experimented with regression analysis to ascertain whether the differential shock effect was attributable in part to intercity wage changes. We found that wage changes explained little, if any, of the differential shock effect. As a result, no wage variables were included in the final regression equation.

## **Conclusion**

This study has provided alternative evidence supporting the significant border effect on relative price volatility. Using a decomposition of relative price volatility different from that considered by Engel & Rogers (1996), this study re-evaluates both the size and the general applicability of the border effect. The decomposition also enables us to quantify the exact contribution of exchange rate volatility to the border effect. It circumvents the collinearity problem that afflicts the regression method typically used in other studies.

Moreover, the decomposition analysis offers potentially useful information about the channels or sources through which the border effect may come from. The border effect is shown to capture both the exchange rate effect and the effect of dissimilar price shocks. Unlike the exchange rate effect, the dissimilar-shock effect has not been well analyzed in the literature. Exploring this non-exchange rate channel may help identify alternative sources of the border effect. Analytically, any micro or macroeconomic factors that can induce significant differential price movement across countries may be a potential contributor. Apart from recognizing the exchange rate contribution, previous studies have focused mainly on microeconomic factors (e.g. geographic distance, trade barriers and costs, and marketing structures) to account for the remaining unexplained border effect. This analysis suggests that macroeconomic shocks can be sufficiently dissimilar across countries to generate extra volatility in cross-border relative prices by inducing differential relative price movement. A pertinent question then is,



how much of the observed border effect can be explained by the cross-country dissimilarity of macroeconomic shocks? This presents an interesting empirical issue for future research.

Overall, the decomposition analysis provides new insights into the estimation, contributing components, and general existence of the border effect. The main results are summed up as follows:

- (1) When examining general cases that allow for cross-country heterogeneity in price volatility, unbiased estimates of the border effect may still be obtained when equal numbers of foreign and home city pairs are sampled from the data.
- (2) When a skewed sample is used without any proper adjustment to correct the heterogeneity bias, the bias can be upward or downward, depending on the specific sample mix from the heterogeneous populations.
- (3) The sufficient conditions for the border effect to occur are: (i) fluctuating exchange rates coupled with sticky prices, and (ii) cross-border relative prices are subject to more dissimilar and less common shocks than are within-country relative prices. Other things being equal, the more (less) significant the dissimilar (common) shocks across countries, the greater the border effect.
- (4) The two conditions for the existence of the border effect are generally valid, suggesting that the border effect on relative price volatility is likely to exist widely in cross-border price data.
- (5) Consistent with sticky prices, exchange rate fluctuations are actually found to contribute to a large majority of the border effect, much higher than the indirect estimate of about 30% reported by the Engel & Rogers study. On average, across goods indices, about 60% of the border effect is explained by exchange rate fluctuations, leaving 40% to be explained by non-exchange rate sources.

### Acknowledgements

The authors are appreciative of all the comments from two anonymous referees. They are grateful to Michael Dooley, Charles Engel, Paul De Grauwe, Ronald McDonald, Michael Melvin, Jan-Egbert Sturm, Shang-Jin Wei, and participants of the CESifo Area Conference and the ASU seminar for additional comments on an earlier version of the manuscript. We would also like to thank John Rogers for providing the Canadian and US data used in the study.

### References

- Cecchetti, S.G., Mark, N.C. & Sonora, R.J. (2002) Price index convergence among United States cities, *International Economic Review*, 43, pp. 1081–1099.
- Ceglowski, J. (2003) The law of one price: intranational evidence for Canada, *Canadian Journal of Economics*, 36, pp. 373–400.

- Cheung, Y.W. & Fujii, E. (2005) Cross-country relative price volatility: effects of market structure, *Review of International Economics*, 14, pp. 836–848.
- Crucini, Mario J., Telmer, Chris I. & Zachariadis, Marios (2005) Understanding European real exchange rates, *American Economic Review*, 95, pp. 724–738.
- Depken, C.A. & Sonora, R.J. (2002) International price volatility: evidence from U.S. and Mexican cities, *North American Journal of Economics and Finance*, 13, pp. 179–193.
- Engel, C. (1993) Real exchange rates and relative prices: an empirical investigation, *Journal of Monetary Economics*, 32, pp. 35–50.
- Engel, C. (1999) Accounting for U.S. real exchange rate changes, *Journal of Political Economy*, 107, pp. 507–538.
- Engel, C. & Rogers, J.H. (1996) How wide is the border? *American Economic Review*, 86, pp. 1112–1125.
- Engel, C. & Rogers, J.H. (2000) Relative price volatility: what role does the border play? In: Gregory D. Hess & Eric van Wincoop (eds) *Intra-national Macroeconomics*, pp. 92–111 (Cambridge University Press).
- Engel, C. & Rogers, J.H. (2001a) Violating the law of one price: should we make a federal case out of it? *Journal of Money, Credit and Banking*, 33, pp. 1–15.
- Engel, C. & Rogers, J.H. (2001b) Deviations from purchasing power parity: causes and welfare costs, *Journal of International Economics*, 55, pp. 29–57.
- Feenstra, R.C. & Kendall, J.D. (1997) Pass-through of exchange rates and purchasing power parity, *Journal of International Economics*, 43, pp. 237–261.
- Gorodnichenko, Y. & Tesar, L. (2005) A re-examination of the border effect. NBER Working Paper No. 11706.
- Parsley, D.C. & Wei, S.-J. (2001) Explaining the border effect: the role of exchange rate variability, shipping costs, and geography, *Journal of International Economics*, 55, pp. 87–105.

### Appendix: Inter- versus Intra-group Pairings from Heterogeneous Populations

To illustrate the generality of our point, let us look at a non-economic example. Suppose that there are two baskets of balls, one carrying red balls (each weighs  $\omega^R$  pounds) and the other carrying blue balls (each weighs  $\omega^B$  pounds, with  $\omega^B < \omega^R$ ). The condition,  $\omega^B < \omega^R$ , indicates that the two ball populations differ in average weight. Two red balls (say,  $R_1$  and  $R_2$ ) and three blue balls (say,  $B_1$ ,  $B_2$ , and  $B_3$ ) are sampled from the baskets, and these balls are paired off two at a time in different possible combinations. The mixed-color sample has six possible pairs:  $(R_1, B_1)$ ,  $(R_1, B_2)$ ,  $(R_1, B_3)$ ,  $(R_2, B_1)$ ,  $(R_2, B_2)$ , and  $(R_2, B_3)$ , while the same-color sample has four possible pairs:  $(R_1, R_2)$ ,  $(B_1, B_2)$ ,  $(B_1, B_3)$ , and  $(B_2, B_3)$ . For mixed-color pairs, the average ball weight is given by  $(\omega^R + \omega^B)/2$ . For same-color pairs, the average weight is given by  $(\omega^R + 3\omega^B)/4$ . The difference in average weight between mixed-color and same-color pairs is given by  $\delta = (\omega^R - \omega^B)/4 > 0$ . This will incorrectly suggest that simply mixing balls of different colors together can add to weight.

If equal numbers of red and blue balls were sampled (e.g. three red and three blue balls), however, the average ball weight for same-color pairs would equal  $(\omega^R + \omega^B)/2$ . For mixed-color pairs, the average weight would still equal  $(\omega^R + \omega^B)/2$ . The difference in average weight between mixed-color and same-color pairs would then be given by  $\delta = 0$ . Hence,

non-skewed samples would correctly show that mixing balls does not alter weight.

In essence, this example shows that skewed samples can produce biased estimates when parings are drawn from heterogeneous populations and that unbiased estimates can be fully secured by simply using non-skewed samples.