Erratum To

Akinwande Atanda
W. Robert Reed

WORKING PAPER
No. 15/2018

Department of Economics and Finance
School of Business
University of Canterbury
Private Bag 4800, Christchurch
New Zealand
Abstract: In a recent paper, Atanda, Menclova and Reed (2018) use country-level panel data from the OECD and conclude that there is little evidence to support the existence of Baumol’s Cost Disease as an explanation for rising health costs. The result is surprising because Hartwig (2008), using similar data, comes to the opposite conclusion. We show that Hartwig tested an incorrect specification of a key hypothesis. When the correct specification is tested, his result vanishes, invalidating his conclusion. This provides a resolution for the conflict between the two studies.

Keywords: Baumol's cost disease, health care expenditures, health care costs, OECD, panel data

JEL Classifications: I11, J30, E24

1 Department of Economics and Finance, University of Canterbury, NEW ZEALAND

† Corresponding author is W. Robert Reed. His email address is bob.reed@canterbury.ac.nz.
Introduction. In a recent paper, Atanda, Menclova and Reed (2018) use country-level panel data from the OECD and conclude that there is little evidence to support the existence of Baumol’s Cost Disease as an explanation for rising health costs. The result is surprising because Hartwig (2008), using similar data, comes to the opposite conclusion. In this erratum, we explain that Hartwig made a fundamental mistake in his estimation that invalidates his conclusion.

Baumol’s Cost Disease. In a series of papers, William J. Baumol (Baumol, 1967; Baumol, 1993; and Baumol and Towse, 1997) argues that relatively non-productive industries are forced to pay higher wages to match compensation in relatively productive industries. This drives wages above marginal product in the non-productive industries. Over time, these wage costs in excess of productivity in the non-productive sector consume ever-increasing shares of GDP, as costs in the non-productive industries are pushed up by productivity increases in the productivity sector. This phenomenon, known as Baumol’s Cost Disease (BCD), is frequently cited as the explanation for why service industries like health, education, and the arts are becomingly increasingly expensive relative to the output they produce.

Hartwig’s Test of BCD. While BCD is consistent with observed increases in health care expenditures, it has been difficult to develop testable hypotheses to confirm its existence. In 2008, Hartwig (2008) published an influential paper\(^1\) in the *Journal of Health Economics* where he claimed to have developed a test for BCD and to have confirmed its existence. While not derived from first principles, the test that Hartwig develops is based around the following idea (Hartwig, page 608): “Baumol’s model states that wage increases in excess of labor productivity growth – averaged across both sectors – drive the rise in health expenditure…”

\(^1\) At the time of this writing, Hartwig (2008) has been cited 76 times in Web of Science.
Hartwig then goes on to state that these wage increases push up costs in a “directly proportional manner”. Accordingly, he sets out to “test this statement empirically”.

Specifically, Hartwig tests the following model:

\[
d \log(HCEPC) = \beta_0 + \beta_1 d \log(WSPE) + \beta_2 d \log(GDPR) + \beta_3 d \log(EMP) + \text{error},
\]

where \(HCEPC\) is health care expenditures per capita, \(WSPE\) is wages and salaries per employee, \(GDPR\) is real GDP, and \(EMP\) is total employment. All the variables are differenced in logs to generate growth rates. He uses country-level data from 19 OECD countries over the years 1971-2003.

His test of BCD consists of two parts. First, he tests whether the variables \(d \log(WSPE), d \log(GDPR),\) and \(d \log(EMP)\) can be combined into a single “Baumol” variable, defined by

\[
(2) \quad Baumol = d \log(WSPE) - d \log(GDPR)
= d \log(WSPE) - d \log(GDPR) + d \log(EMP)
\]

Next, he tests whether the coefficient on the Baumol variable equals 1 in the specification below:

\[
(3) \quad d \log(HCEPC) = \alpha_0 + \alpha_1 Baumol + \text{error}.
\]

After performing this two-part test, Hartwig concludes that he has found evidence of the existence of BCD:

“Proceeding in a general-to-specific manner, we first estimate the influence of these three variables separately in order to test whether the restriction of summing them together to one variable is legitimate. The estimation period covers the years from 1971 to 2003. Table 1 shows that the three coefficients are statistically different from zero with signs as expected. For all three estimations, Wald test results (shown in the bottom line of the table) fail to reject the hypothesis that \(C(1)+C(2)−C(3) = 0\) so that we can legitimately combine the three variables into one. Table 2 shows our results for the ‘Baumol variable’. We find that Baumol’s model of ‘unbalanced growth’ is strongly supported by the data. In all three specifications, the coefficient of the difference between nominal wage and productivity growth rates is statistically different from zero. As predicted by Baumol’s theory, the value of the coefficient is close to one.
Again, the Wald test fails to reject the hypothesis that the coefficients are in fact equal to one” (page 610).

From the above, it is clear that a crucial component of Hartwig’s test for BCD is testing “whether the restriction of summing them [the three variables] together to one variable is legitimate”. Notice that to implement his test, he tests whether “\(C(1)+C(2)-C(3)=0\)”. Translated in the context of Equation (1), this equates to testing \(H_0: \beta_1 + \beta_2 - \beta_3 = 0\). This is what Hartwig tests, and on the basis of failing to reject this hypothesis, he combines the three variables into a single “Baumol variable”, estimates Equation (3) above, and fails to reject that the coefficient on the Baumol variable, \(\alpha_1\), equals 1.

Replication. The problem is that Hartwig tests the wrong hypothesis in the first step. The correct test for determining if the three variables, \(d \log(WSPE), d \log(GDPR),\) and \(d \log(EMP)\), can be combined into a single variable via Equation (2) is the joint hypothesis, \(H_0: \beta_1 = -\beta_2\) and \(H_0: \beta_1 = \beta_3\), not \(H_0: \beta_1 + \beta_2 - \beta_3 = 0\). When the correct hypothesis is tested, it is strongly rejected, so that we conclude that it is not legitimate to combine the three variables into a single Baumol variable. It follows that the the second part of Hartwig’s test for BCD is invalid.

Table 1 reports the results of replicating Table 1 in Hartwig (2008). Our results exactly match his for each of the three estimation procedures: OLS, Cross-section Random Effects, and Time period Random Effects. Note particularly the results of the Wald tests at the bottom of the table. For the OLS procedure, after estimating Equation (1), Hartwig tests \(H_0: \beta_1 + \beta_2 - \beta_3 = 0\). He obtains a sample \(F\) statistic of 0.817. The associated p-value is 0.366. Our results exactly match his. However, this is the wrong test! When we test the correct (joint) hypothesis, \(H_0: \beta_1 = -\beta_2; H_0: \beta_1 = \beta_3\), we obtain a sample \(F\) value of 34.068 which has an

---

2 An accompanying Eviews workfile is provided with this manuscript that allows one to replicate all the results in Table 1.
associated p-value of 0.000. In other words, we find very strong evidence that it is not valid to combine the three variables into a single Baumol variable. Similar results follow for the other two specifications in Table 1.

**Conclusion.** In a recent paper, Atanda, Menclova, and Reed (2018) examined annual health care expenditures for 28 OECD countries over the years 1995–2016. They found no evidence to support the existence of Baumol’s Cost Disease (BCD) as a determinant of rising health costs. This conflicts with Hartwig (2008), who used similar data over a similar time period. While the studies are not identical -- for example, they formulate different hypotheses to test BCD -- it is still troubling that they arrive at opposite results. This note provides an explanation for this result.

In particular, we find that Hartwig tested the wrong thing. He proposed a test for BCD, but when he went to implement that test, he formulated an incorrect hypothesis. We show that when one tests the correct hypothesis, Hartwig’s test for BCD fails. This provides a resolution for the conflict between the two studies.
REFERENCES


TABLE 1
Replication of Hartwig’s Table 1: Results for Growth Rate Equations

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) Cross-section R.E</th>
<th>(3) Time period R.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d \log(WSPE))</td>
<td>1.066* (28.557)</td>
<td>1.064* (27.561)</td>
<td>1.059* (27.155)</td>
</tr>
<tr>
<td>(d \log(GDPR))</td>
<td>-0.339* (-3.951)</td>
<td>-0.351* (-4.049)</td>
<td>-0.308* (-3.571)</td>
</tr>
<tr>
<td>(d \log(EMP))</td>
<td>0.601* (7.377)</td>
<td>0.599* (7.331)</td>
<td>0.588* (7.511)</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>507</td>
<td>507</td>
<td>507</td>
</tr>
<tr>
<td>(R^2) (adj.)</td>
<td>0.810</td>
<td>0.798</td>
<td>0.799</td>
</tr>
<tr>
<td>Stand. Err. of regr.</td>
<td>0.032</td>
<td>0.031</td>
<td>0.031</td>
</tr>
<tr>
<td>D-W</td>
<td>1.830</td>
<td>1.852</td>
<td>1.827</td>
</tr>
</tbody>
</table>

Wald test \(F\)-stat (prob.)

\(H_0: \beta_1 + \beta_2 - \beta_3 = 0\)
- (1) OLS: 0.817 (0.366)
- (2) Cross-section R.E: 0.648 (0.421)
- (3) Time period R.E: 1.318 (0.252)

\(H_0: \beta_1 = -\beta_2; H_0: \beta_1 = \beta_3\)
- (1) OLS: 34.068 (0.000)
- (2) Cross-section R.E: 31.557 (0.000)
- (3) Time period R.E: 33.500 (0.000)
NOTE: HCE is health care expenditures per capita, WSPE is wages and salaries per employee, GDPR is real GDP, and EMP is total employment. All variables are differenced in logs to generate growth rates. The estimated equation is given in the top of the table. Coefficient estimates for the three explanatory variables are reported in the first three rows, with t-stats reported in parentheses. The estimate for the constant term is not reported. The data consist of annual, country-level observation from 19 OECD countries over the years 1971-2003. “OLS”, “Cross-section Random Effects”, and “Time period Random Effects” refer to the following three “pool estimation” procedures in Eviews: pooled OLS, Random Effects with random effects for cross-sections, and Random Effects with random effects for time periods. In all three cases, White’s robust estimator for cross-sectional dependence is used to estimate standard errors. A “*” indicates the coefficient is statistically significant at the 1 percent level.