

Erratum

Jeffrey R. Campbell* Hugo A. Hopenhayn†

June, 2006

In the March 2005 edition of the *Journal of Industrial Economics*, we published an article entitled “Market Size Matters” (Volume 53, Number 1, pages 1–25). The original paper asserted that retail establishments in large markets have greater average sales and employment. The paper supported this assertion by showing that regressions of retailers’ average sizes on market sizes yielded positive and significant coefficients for most of the retail industries under consideration, even after controlling for differences between markets’ factor prices and demographics. The paper reported estimates based on three estimation techniques, ordinary least squares, instrumental variables, and nonparametric density-weighted average derivative estimation. While recently extending that paper’s results, we discovered a coding error in the Gauss procedure we wrote to calculate estimates of density-weighted average derivatives.¹ We have corrected this error, and we have found that the correctly calculated estimates quantitatively support the article’s original conclusion. The purpose of this erratum is to present corrected tables for the paper.

Following our discovery of the programming error, we examined all of the paper’s results for possible errors. In the process, we found some relatively inconsequential errors in the original tables due to incorrect transcription of original program output. The corrected tables are reported as Tables I through VI below. These table numbers correspond to those in the original article.

*Federal Reserve Bank of Chicago and NBER. mail: Federal Reserve Bank of Chicago, 230 South LaSalle St., Chicago, Illinois 60604–1413. e-mail: jcampbell@frbchi.org

†University of California, Los Angeles; and Universidad Torcuato Di Tella. mail: Department of Economics, UCLA, Box 951477, Los Angeles, CA 90095–1477. e-mail: hopen@econ.ucla.edu

¹The specific error involved the implementation of the following operation. Given two identically sized column vectors x and y , sum the elements of x for which the corresponding elements of y weakly exceed zero. In Matlab, this can be done with `sum(x(y>=0))`. If all elements of y are strictly negative, then `x(y>=0)` returns an empty matrix, and the sum of the empty matrix’s elements is zero. To conduct this operation in Gauss, we calculated the indices of y ’s non-negative elements with `i=indexcat(y.>=0,1)`. We then calculated the sum of the corresponding elements of x with `sumc(x[i])`. If all of the elements of y are non-positive, then `indexcat` sets the variable `i` equal to a missing value code. In Gauss, indexing a vector with a missing value code returns the entire vector. Hence, if all of the elements of y are non-positive, the Gauss code we used returned the sum of *all* the elements of x . Our program would have worked correctly if Gauss interpreted a matrix indexed with a missing value as Matlab does. However, we failed to account for this aspect of the Gauss language. The direct consequence of this error was the incorrect calculation of non-parametric density estimates.

The differences between Tables I, II, and III in this erratum and their versions in the original article reflect only transcription errors.² The next two tables report the regression estimates that serve as the primary evidence in favor of the paper’s conclusion. Table IV contains estimates from regressions of establishments’ average sales on market size and other control variables, while Table V contains estimates from analogous regressions using establishments’ average employment instead. All of the estimates are coefficients on market size. Both Tables’ fourth columns report density-weighted average derivative estimates. In the original paper, the estimates in the fourth column of Table IV were all positive and statistically significant at the 5% level for ten of the thirteen industries. Correcting the estimates’ calculation eliminated the statistical significance at the 5% level for two industries, Grocery Stores and Shoe Stores, and raised the statistical significance past the 5% level for one industry, Refreshment Places. The statistical inferences for the other industries remain unchanged, although the correction substantially reduced some coefficients’ estimated magnitudes.

The correction impacts the fourth column of Table V more broadly. In the original version, nine of the thirteen industries had positive coefficients that were statistically significant at the 5% level. The corrected estimates for five of these industries remain statistically significant, and previously insignificant estimates from two industries (New and Used Car Dealers and Radio/TV/Computer/Music Stores) become statistically significant. One industry (Auto and Home Supply Stores) had a counterintuitive negative and statistically significant coefficient. For that industry, the corrected estimate now equals zero.

Table VI reports density-weighted average derivative estimates from regressions of the employment *c.d.f.* at three predefined levels – nine, nineteen, and forty-nine employees, on *MSA* population. The estimates from the first three columns come from specifications that included the control variables listed in Table II, and the final three columns’ estimates used no controls. The results reported in the original paper indicated that the *dispersion* of establishment sizes changed systematically with market size for three industries. The corrected results lead us to a different conclusion. With the exception of Women’s Clothing and Specialty Stores, most of the estimated coefficients are statistically insignificant. This finding, together with the result that average employment increases with market size for most of these industries, implies that the effects of market size operate primarily on firms with fifty or more employees.

To facilitate further replication of our results, we have created a replication file containing Matlab and L^AT_EX code for automatically generating this erratum. It is available at

<http://www.nber.org/~jrc/marketsizematters>.

²We have verified that all differences between the tables reported in this erratum and those in the original article are due to either the specific programming error mentioned above or due to transcription errors. For this, we conducted two independent reviews of the original statistical output.

Table I: Summary Statistics from *Census of Retail Trade* and *County Business Patterns* Data⁽ⁱ⁾



Industry	SIC Code(s) ⁽ⁱⁱ⁾	Average Sales ^(iii,iv,v)			Average Employment ^(iii,iv)			Employment <i>c.d.f.</i> ^(vi)		
		<i>Q</i> (1)	Median	<i>Q</i> (3)	<i>Q</i> (1)	Median	<i>Q</i> (3)	<i>F</i> (9)	<i>F</i> (19)	<i>F</i> (49)
Building Materials and Supplies	521,3	1901	2236	2603	11.4	13.3	15.9	0.66	0.82	0.94
Grocery Stores	541	2354	2821	3517	18.5	22.9	28.0	0.62	0.75	0.85
New and Used Car Dealers	551	11074	13984	18284	32.5	38.8	47.5	0.22	0.35	0.73
Auto and Home Supply Stores	553	612	701	795	6.1	6.9	7.8	0.75	0.94	1.00
Gasoline Service Stations	554	1106	1287	1422	5.9	6.6	7.6	0.81	0.97	0.99
Women's Clothing and Specialty Stores	562,3	431	493	563	6.6	7.5	8.4	0.70	0.94	1.00
Shoe Stores	566	401	444	500	4.6	5.0	5.5	0.90	0.99	1.00
Furniture Stores	5712	746	911	1074	6.3	7.4	8.8	0.74	0.92	0.99
Homefurnishings Stores	5713,4,9	474	559	642	4.6	5.3	6.2	0.85	0.97	1.00
Radio/TV/Computer/Music Stores	573	641	825	990	5.4	6.3	7.6	0.83	0.95	0.99
Restaurants ^(vii)	5812 ^(viii)	433	494	555	17.0	19.3	21.6	0.42	0.64	0.91
Refreshment Places	5812 ^(viii)	462	502	547	16.5	18.3	20.0	-	-	-
Drug and Proprietary Stores	591	1275	1501	1881	10.5	12.3	14.7	0.51	0.82	0.97

Notes: (i) For each of these industries, the fifth percentile of the number of establishments across all *MSAs* equals or exceeds 10. (ii) When multiple SIC codes are given, the industry is defined as the union of those industries. (iii) The headings *Q* (1) and *Q* (3) refer to the first and third sample quartiles. (iv) The entries in each column are the sample quartiles, across *MSAs*, of establishments' average sales and employment for that industry. (v) Average sales is reported in thousands of 1992 dollars. (vi) The entries in each column are the sample averages across *MSAs* of the empirical *c.d.f.* of employment evaluated at 9, 19, and 49 employees. All entries are reported to two significant digits. The underlying estimates are all strictly less than one. (vii) The reported average *c.d.f.* reported for Restaurants is for its parent industry, Eating Places. (viii) These industries are subsets of SIC 5812, Eating Places. Restaurants are those establishments that provide table service. See the text for further details.

Table II: Independent Variables Used in the Regressions

Variable	Description	Source ⁽ⁱ⁾	$Q(1)$	Median	$Q(3)$	Correlation ⁽ⁱⁱ⁾
Population	Total MSA Residents	CCDB	136734	254861	471837	1.00
Retail Wage	First Quarter Retail Payroll/ March Employment	CRT	2484	2586	2724	0.39
Commercial Rent ^(iv)	Median Rent per Square Foot for Strip Malls	CH ^(v)	7.00	8.00	9.50	0.33
Advertising Cost ⁽ⁱⁱⁱ⁾	Cost of Standard Ad in Sunday Newspaper	CH ^(v)	0.43	0.51	0.58	-0.40
Income	Per Capita Personal Income	BEA	17376	18668	20407	0.37
Percent Black	% of Population that is Black	CCDB	2.73	7.11	16.23	0.09
Percent College	% of Population over 25 with a College Degree	CCDB	14.83	18.04	21.62	0.10
Vehicle Ownership	Vehicles per Household	CCDB	1.66	1.71	1.80	-0.16

Notes: (i) CCDB is the 1994 County and City Data Book, CRT is the 1992 Census of Retail Trade, BEA is the Bureau of Economic Analysis Regional Accounts File, and CH denotes the authors' calculations. (ii) These correlations are calculated using the logarithm of population and, depending on how it enters our regressions, either the logarithm or the level of the indicated variable. (iii) In 1992 dollars. (iv) In 1992 dollars per square foot. (v) Our observations of rent per square foot for strip malls comes from the 1993 Shopping Center Directory. (vi) Our observations of Sunday newspaper advertising rates and circulation come from the 1992 Editor and Publisher International Yearbook. See the text and ?? for further details regarding the data's construction.

Table III: OLS Estimation Results

	Estimates for Women's Clothing ⁽ⁱ⁾		+/- Table for all Industries ⁽ⁱⁱ⁾	
	Average Sales	Average Employment	Average Sales	Average Employment
Population	0.10 ^{***} (0.01)	0.06 ^{***} (0.02)	7/0	6/0
Retail Wage	-0.08 (0.21)	-0.46 ^{***} (0.17)	5/0	2/4
Commercial Rent	-0.05 (0.06)	-0.03 (0.05)	0/0	0/0
Advertising Cost	-0.02 (0.05)	-0.02 (0.05)	0/1	0/1
Income	0.19 (0.16)	0.33 ^{***} (0.11)	3/1	5/0
Percent Black ⁽ⁱⁱⁱ⁾	-0.03 (0.12)	-0.14 (0.09)	3/6	2/4
Percent College ⁽ⁱⁱⁱ⁾	0.55 ^{**} (0.24)	0.54 ^{**} (0.25)	9/0	10/0
Vehicle Ownership	-0.48 ^{***} (0.12)	-0.38 ^{***} (0.11)	3/3	1/1
R^2	0.32	0.26		

Note: (i) Heteroskedasticity consistent White standard errors appear below each estimate in parentheses. The superscripts *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. (ii) Each cell's first element gives the number of retail trade industry regressions in which the corresponding t -statistic is greater than or equal to 1.96, and each cell's second element gives the number of such regressions in which the t -statistic is less than or equal to -1.96. (iii) For comparability, the estimated coefficients on these variables and their standard errors are multiplied by 100. See the text for further details.

Table IV: Market Size Effects on Average Sales^(i,ii)

	Population				Density
	OLS	No Controls	IV	PSS	OLS
Building Materials and Supplies	0.03 (0.02)	0.07*** (0.02)	0.02 (0.03)	0.05* (0.03)	0.06*** (0.02)
Grocery Stores	0.00 (0.02)	0.01 (0.02)	0.01 (0.03)	-0.01 (0.03)	-0.00 (0.02)
New and Used Car Dealers	0.08*** (0.03)	0.18*** (0.02)	0.15*** (0.05)	0.07*** (0.03)	0.06** (0.02)
Auto and Home Supply Stores	-0.01 (0.02)	0.01 (0.02)	-0.01 (0.02)	-0.03 (0.02)	0.01 (0.02)
Gasoline Service Stations	0.05*** (0.02)	0.10*** (0.02)	0.02 (0.03)	0.05** (0.02)	0.02 (0.02)
Women's Clothing and Specialty Stores	0.10*** (0.01)	0.12*** (0.01)	0.10*** (0.03)	0.09*** (0.02)	0.09*** (0.02)
Shoe Stores	0.02 (0.01)	0.05*** (0.02)	0.00 (0.02)	0.02 (0.01)	0.05*** (0.02)
Furniture Stores	0.11*** (0.03)	0.12*** (0.02)	0.17*** (0.05)	0.11*** (0.04)	0.12*** (0.03)
Homefurnishings Stores	0.05** (0.02)	0.08*** (0.02)	0.08** (0.04)	0.05** (0.02)	0.03* (0.02)
Radio/TV/Computer/Music Stores	0.16*** (0.02)	0.17*** (0.02)	0.19*** (0.03)	0.14*** (0.03)	0.09*** (0.03)
Restaurants	0.05*** (0.02)	0.09*** (0.01)	0.05*** (0.02)	0.05*** (0.02)	0.04** (0.02)
Refreshment Places	0.02* (0.01)	0.01 (0.01)	0.01 (0.01)	0.03** (0.01)	0.04*** (0.01)
Drug and Proprietary Stores	0.03 (0.02)	0.10*** (0.02)	-0.04 (0.04)	0.03 (0.02)	0.05** (0.02)

Notes: (i) The table's entries are estimated coefficients on the logarithm of market size from the industry-specific regressions described in the text. Heteroskedasticity-consistent standard errors appear in parentheses.

(ii) The superscripts *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. See the text for further details.

Table V: Market Size Effects on Average Employment^(i,ii)

	Population				Density	Total Sales
	OLS	No Controls	IV	PSS	OLS	OLS
Building Materials and Supplies	0.06 ^{***} (0.02)	0.08 ^{***} (0.02)	0.05 [*] (0.03)	0.07 ^{***} (0.03)	0.08 ^{***} (0.02)	0.14 ^{***} (0.03)
Grocery Stores	0.01 (0.03)	0.01 (0.02)	0.03 (0.03)	0.01 (0.03)	0.01 (0.02)	0.01 (0.03)
New and Used Car Dealers	0.06 ^{***} (0.02)	0.13 ^{***} (0.02)	0.11 ^{***} (0.04)	0.07 ^{***} (0.02)	0.05 ^{**} (0.02)	0.08 ^{***} (0.02)
Auto and Home Supply Stores	0.01 (0.02)	0.01 (0.01)	0.03 (0.02)	0.00 (0.02)	0.03 [*] (0.02)	0.05 ^{***} (0.02)
Gasoline Service Stations	0.02 (0.02)	0.04 ^{**} (0.02)	-0.01 (0.03)	0.02 (0.02)	0.01 (0.02)	0.06 ^{***} (0.02)
Women's Clothing and Specialty Stores	0.06 ^{***} (0.02)	0.08 ^{***} (0.01)	0.09 ^{***} (0.03)	0.05 ^{***} (0.02)	0.06 ^{***} (0.02)	0.09 ^{***} (0.01)
Shoe Stores	0.02 (0.01)	0.01 (0.01)	0.01 (0.02)	0.02 (0.01)	0.06 ^{***} (0.01)	0.06 ^{***} (0.02)
Furniture Stores	0.05 [*] (0.03)	0.04 ^{**} (0.02)	0.10 ^{***} (0.04)	0.04 (0.03)	0.08 ^{***} (0.03)	0.13 ^{***} (0.03)
Homefurnishings Stores	0.05 ^{**} (0.02)	0.08 ^{***} (0.02)	0.09 ^{**} (0.04)	0.05 ^{**} (0.02)	0.05 ^{**} (0.02)	0.10 ^{***} (0.02)
Radio/TV/Computer/Music Stores	0.07 ^{***} (0.02)	0.08 ^{***} (0.02)	0.09 ^{***} (0.03)	0.06 ^{**} (0.02)	0.04 (0.03)	0.13 ^{***} (0.02)
Restaurants	0.05 ^{***} (0.02)	0.05 ^{***} (0.01)	0.07 ^{***} (0.02)	0.06 ^{***} (0.02)	0.05 ^{***} (0.02)	0.07 ^{***} (0.02)
Refreshment Places	0.02 [*] (0.01)	-0.01 (0.01)	0.03 [*] (0.02)	0.03 ^{**} (0.01)	0.02 ^{**} (0.01)	0.04 ^{***} (0.01)
Drug and Proprietary Stores	0.03 (0.02)	0.07 ^{***} (0.02)	-0.02 (0.04)	0.02 (0.02)	0.04 ^{**} (0.02)	0.07 ^{***} (0.02)

Notes: (i) The table's entries are estimated coefficients on the logarithm of market size from the industry-specific regressions described in the text. Heteroskedasticity-consistent standard errors appear in parentheses. (ii) The superscripts *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. See the text for further details.

Table VI: Market Size Effects on Employment's *c.d.f.*^(i,ii,iii)

Industry	Controls Included			No Controls		
	<i>F</i> (9)	<i>F</i> (19)	<i>F</i> (49)	<i>F</i> (9)	<i>F</i> (19)	<i>F</i> (49)
Building Materials and Supplies	-0.17 (0.74)	0.00 (0.55)	-0.55 (0.35)	0.51 (0.64)	-0.23 (0.56)	-1.29*** (0.28)
Grocery Stores	0.24 (0.98)	0.46 (0.66)	-0.10 (0.52)	1.19 (0.76)	0.73 (0.49)	-0.12 (0.42)
New and Used Car Dealers	0.74 (0.95)	0.53 (1.08)	-1.72 (1.20)	1.05 (0.72)	-0.53 (0.90)	-4.98*** (0.92)
Auto and Home Supply Stores	0.31 (0.88)	0.21 (0.38)	0.08 (0.10)	-0.38 (0.65)	0.11 (0.30)	0.08 (0.09)
Gasoline Service Stations	-0.85 (0.83)	0.06 (0.27)	-0.07 (0.07)	-1.76** (0.73)	-0.17 (0.22)	0.02 (0.07)
Women's Clothing and Specialty Stores	-1.93*** (0.72)	-1.20*** (0.38)	-0.17** (0.07)	-1.64*** (0.57)	-1.61*** (0.26)	-0.24*** (0.06)
Shoe Stores	1.00* (0.54)	-0.03 (0.17)	-0.01 (0.03)	0.69 (0.51)	-0.06 (0.20)	-0.03 (0.05)
Furniture Stores	2.12* (1.25)	0.06 (0.65)	-0.50* (0.29)	0.97 (0.80)	-0.33 (0.48)	-0.49*** (0.17)
Homefurnishings Stores	-0.73 (0.76)	-0.79** (0.35)	-0.09 (0.09)	-1.56** (0.61)	-0.74** (0.32)	-0.16** (0.06)
Radio/TV/Computer/Music Stores	-0.98 (0.74)	-0.58 (0.41)	-0.26* (0.14)	-1.07* (0.56)	-0.87*** (0.33)	-0.33*** (0.12)
Eating and Drinking Places	0.19 (0.64)	-0.67 (0.51)	-0.32 (0.26)	1.51*** (0.53)	0.37 (0.44)	-0.33 (0.21)
Drug and Proprietary Stores	-0.76 (1.23)	0.13 (0.90)	0.11 (0.30)	-2.82*** (1.00)	-2.38*** (0.79)	-0.44* (0.27)

Notes: (i) The table's entries are estimated density-weighted average derivatives, expressed in probability points, of the indicated variable with respect to the logarithm of *MSA* Population. Heteroskedasticity-consistent standard errors appear in parentheses. (ii) The superscripts *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. (iii) In the column headings, *F* (9), *F* (19) and *F* (49) refer to the empirical *c.d.f.* of the distribution of employment across an *MSA*'s establishments. "Controls Included" and "No Controls" refer to regressions with and without the control variables listed in Table II.