Equivalence Scales

An Update and an Extension

Steven F. Koch, PhD

University of Pretoria
Department of Economics

June 2017

Outline

South African Estimates

- Woolard & Leibbrandt (1999) and Woolard (2002)
  - $E = (a + \beta_2 k)_1^{\beta}$
  - African households only
  - 1993 & 1995 data
  - Engel method
  - $\hat{\beta}_2 \approx 0.5$ and $\hat{\beta}_1 \approx 0.9$
- Most other studies no scale estimated
  - Adjust with rule-of-thumb from May, Carter & Poset (1995)
  - May, Budlender, Mokate, Rogerson & Stavrou (1995) and Meth & Dias (2004)
- Three exceptions
  - Yatchew et al. (2003) semiparametric
  - Anonymous (2016) linear
  - Koch (2017) semiparametric

Outline

Engel Overview

Engel's Method

- Following Engel's (1857) conceptualization
- x is total expenditure
- n is household size
- n_j is proportion of people in 'group' j
- Z are other controls
- w is food share of budget

$$w = \beta_0 + \beta_1 \ln\left(\frac{x}{n}\right) + \sum_j \gamma_j n_j + Z\delta. \tag{1}$$

- Estimates can be used to calculate an equivalance
- Ignore Z for simplicity

Finding the Scale

- ‘Equalize’ food expenditure shares

$$
\begin{aligned}
w^a &= w^r \\
\beta_0 + \beta_1 \ln\left(\frac{x^a}{n^a}\right) + \sum_j \gamma_j n_j^a &= \beta_0 + \beta_1 \ln\left(\frac{x^r}{n^r}\right) + \sum_j \gamma_j n_j^r \qquad (2)
\end{aligned}
$$

- Solve for the ratio of expenditure
- Rearranging terms

The Engel Scale

- Once completed..

$$E \equiv \frac{x^a}{x^r} = \frac{n^a}{n^r} \exp\left(\frac{\sum_j \gamma_j \left(n_j^r - n_j^a \right)}{\beta_1} \right) \tag{3}$$

- Thus, estimate (1)
- Plug into (3)
- Bootstrap for standard errors

Base Independence

Base Independence: The Idea

- Blundell & Lewbel (1991), extending Pollak & Wales (1979)
  - Different preferences will give same demand curves
  - Eq Scales not identified from demand curves
  - But, cost of living index is estimable
  - Can recover relative CoL
- Equivalence scales are independent of base utility
  - Blackorby & Donaldson (1993) provide a different interpretation
  - Monotonic transformation of utility cannot include demographic structure
  - Income-ratio comparability
  - But, this means Working-Leser shares ‘fail’

Base Independence Applied

- Optimal result: Indirect utility

$$V(p, x, z) = V\left(p, \frac{x}{\Delta(p, x, z)}, z^r\right) \tag{4}$$

- From basic micro theory

$$w_j(p, x, z) = -\frac{\partial V / \partial \ln p_j}{\partial V / \partial \ln x} = -\frac{\partial V / \partial p_j}{\partial V / \partial x} \times \frac{p_j}{x} \tag{5}$$

- This is a semilog derivative

Base Independence Applied: The Numerator

- The semilog derivative

$$w_j(p, x^r, z^r) = -\frac{\partial V/\partial \ln p_j}{\partial V/\partial \ln x^r} \tag{6}$$

- The numerator

$$\begin{aligned} -\frac{\partial V}{\partial p_j} \times \frac{p_j}{x} &= \left[-V_p - V_x \frac{x}{\Delta^2}\left(-\frac{\partial \Delta}{\partial p}\right)\right]\frac{p}{x} \\ &= -V_p \times \frac{p}{x} + \frac{V_x}{\Delta}\left(\frac{\partial \Delta}{\partial p} \times \frac{p}{\Delta}\right) \\ &= -\frac{V_p p}{x} + \frac{V_x}{\Delta}\eta_{\Delta p} \end{aligned} \tag{7}$$

Base Independence Applied: The Denominator

- The semilog derivative

$$w_j(p, x^r, z^r) = -\frac{\partial V / \partial \ln p_j}{\partial V / \partial \ln x^r} \tag{8}$$

- The denominator

$$\begin{aligned} \frac{\partial V}{\partial x} &= \frac{V_x}{\Delta} + V_x \frac{x}{\Delta^2}\left(-\frac{\partial \Delta}{\partial x}\right) \\ &= \frac{V_x}{\Delta}\left(1 - \frac{\partial \Delta}{\partial x}\frac{x}{\Delta}\right) \\ &= \frac{V_x}{\Delta}\left(1 - \eta_{\Delta x}\right) \end{aligned} \tag{9}$$

Base Independence Applied: The Result

- The solution

$$\begin{aligned}
w_j(p,x,z) &= -\left(\frac{V_p p}{x} + \frac{V_x}{\Delta}\eta_{\Delta p}\right)\left(\frac{\Delta}{V_x(1-\eta_{\Delta x})}\right) \\
&= \left(-\frac{V_p p}{x}\frac{\Delta}{V_x(1-\eta_{\Delta x})}\right) + \left(\frac{V_x \eta_{\Delta p}}{\Delta}\frac{\Delta}{V_x(1-\eta_{\Delta x})}\right) \\
&= \left[\frac{1}{1-\eta_{\Delta x}}\right]\left(-\frac{V_p}{V_x}\frac{p\Delta}{x}\right) + \frac{\eta_{\Delta p}}{1-\eta_{\Delta x}} \\
&= \frac{1}{1-\eta_{\Delta x}}\left[\left(-\frac{V_p}{V_x}\frac{p}{x/\Delta}\right) + \eta_{\Delta p}\right]
\end{aligned} \tag{10}$$

Base Independence Applied: The Simplification

- The first simplification

$$w_j(p, x, z) = \frac{w_j(p, x^r, z^r) + \eta_{\Delta p}}{1 - \eta_{\Delta x}} \tag{11}$$

- With Base Independence

$$w_j(x, z) = w_j(x^r, z^r) + \eta_p \tag{12}$$

- Estimation is not obvious

Outline

Semiparametric Methods

An Index Model

- Yatchew et al. (2003) provide a succinct version
- They also estimate for South Africa (a different model)
- Consider the minor generalisation of equation (**??**)

provide a number of useful improvements in

$$y_b = f_b\left(p, x_b\right) = f_a\left(p, \frac{x_b}{\Delta_b(p)}\right) + \eta_b(p) \tag{13}$$

- This can be placed into an index framework

$$y = f\left(\ln x - z\delta\right) + z\eta + \varepsilon \tag{14}$$

The Semiparametric Models: Model 1

- Rearranging Yatchew et al. (2003)
- Version 1

$$
\begin{aligned}
y &= f\left(\ln x - z\delta\right) + z\eta + \varepsilon \\
z\delta &= \theta \ln(a + k) \\
z\eta &= (a + k)\eta \\
y &= f\left(\ln x - \theta \ln[a + k]\right) + [a + k]\eta + \varepsilon
\end{aligned}
\tag{15}
$$

- Now, we just need θ and η
- We undertake grid search
- Employ Robinson (1988) Double Residual Method
- Some minor tweaks to Yatchew et al. (2003)
- *a* nand *k* are adults and children

The Semiparametric Models: Model 2

- Rearranging Yatchew et al. (2003)
- Version 2

$$\begin{aligned} y &= f\left(\ln x - z\delta\right) + z\eta + \varepsilon \\ z\delta &= \beta_2 \ln(a + \beta_1 k) \\ z\eta &= \eta_1 a + \eta_2 k \\ y &= f\left(\ln x - \beta_2 \ln[a + \beta_1 k]\right) + \eta_1 a + \eta_2 k + \varepsilon \end{aligned} \tag{16}$$

- Now, we just need β_1, β_2, η_1 and η_2
- We undertake grid search here, too
- Employ Robinson (1988) Double Residual Method
- Some minor tweaks to Yatchew et al. (2003)
- *a* nand *k* are adults and children

The Semiparametric Models: Model 3

- Leave as is

$$y = f(\ln x - z\delta) + z\eta + \varepsilon \tag{17}$$

- Now, need a series of δs and ηs
- We undertake grid search
- Again Robinson (1988) Double Residual Method
- Some minor tweaks to Yatchew et al. (2003)
- zs are dummy variables for each household structure

Partial Linear Model

- Consider

$$y = f(X) + Z\beta + u \tag{18}$$

- Take expections with respect to Z

$$E[y|X] = f(X) + E[Z|X]\beta \tag{19}$$

- Subtract

$$\begin{aligned} y - E[y|X] &= (Z - E[Z|X])\beta \\ \tilde{y} &= \tilde{Z}\beta \end{aligned} \tag{20}$$

Estimation in Practice

- ‘Revised’ OLS, Robinson (1988)
- But need estimates for $E[y|X]$ and $E[Z|X]$
- Estimate nonparametrically, separately
- Use Hayfield & Racine (2008) in R Core Team (2016)

Partially Linear Index Model

- Consider, where δ might also be a function

$$y = f(X - \delta) + Z\eta + u \tag{21}$$

- Grid search: specify value of δ
- Follow double-residual method on $(X - \delta)$
- Repeat
- Optimum based on minimum sum of squares
- Variance of η: covariance matrix from ‘revised’ OLS
- Variance of δ requires squared derivative of f

$$\begin{aligned} V &= E[f'(X-\delta)^2[X - E[X|X-\delta]][X - E[X|X-\delta]]|X-\delta] \\ V(\delta) &= \sigma_u^2 V^{-1} \end{aligned} \tag{22}$$

Outline

The Data

2010 Income and Expenditure Survey

- Primarily used for CPI
- Contain needed data
  - Food expenditure
  - Total expenditure
  - Household structure
  - Further breakdown by race
- Expenditures follow COICOP
  - Classification of Individual Consumption According to Purpose
  - Initial switch in 2005-06
  - Health is 06

Descriptive Statistics

Table: Descriptive Statistics of 2010 IES Data

	All HH	Black HH	Coloured HH	White HH
Household Size	3.75	3.79	3.84	2.67
Food Expenditure	951.72	831.98	1272.54	1623.81
Food Share	0.25	0.27	0.25	0.09
Total HH Expenditure	6630.26	4585.65	7766.59	23048.87

Descriptive statistics.

A Look at Base Independence: Sort of

Single Adult Black Households

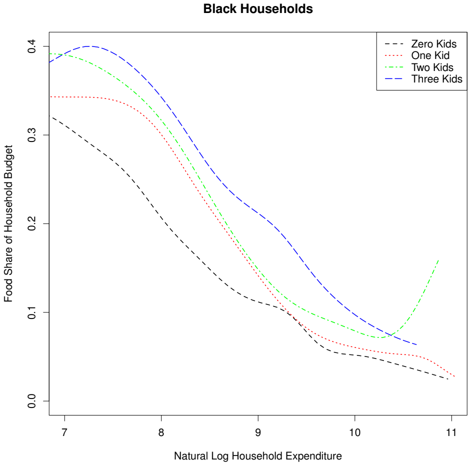

Figure: Computed bandwidths for zero children, one child, two children and three children are 0.209, 0.282, 0.369 and 0.314.

Two Adult Black Households

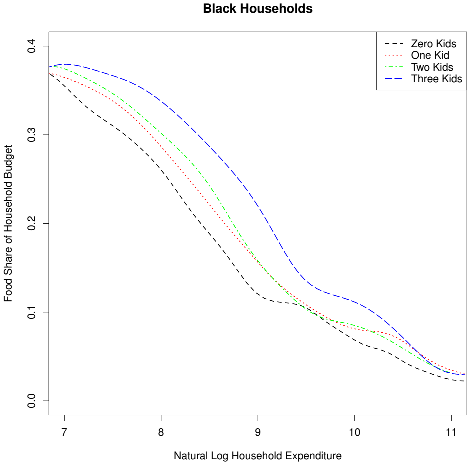

Figure: Computed bandwidths for zero children, one child, two children and three children are 0.179, 0.233, 0.239 and 0.282.

Outline

Semiparametric Estimates

Estimates: Models 1 and 2

Table: Parameter Estimates from Semiparametric Models

	$\hat{\theta}$ (s.e.)	$\hat{\eta}$ (s.e.)	$\hat{\beta}_1$ (s.e.)	$\hat{\beta}_2$ (s.e.)	$\hat{\eta}_1$ (s.e.)	$\hat{\eta}_2$ (s.e.)
All HH SP	0.4275	0.0040	1.0000	0.4375	0.0000	0.0081
N= 24206	(0.024)	(0.001)	(0.031)	(0.048)	(0.001)	(0.001)
Black HH SP	0.3825	0.0047	1.0000	0.3925	-0.0004	0.0098
N= 19143	(0.026)	(0.001)	(0.035)	(0.051)	(0.002)	(0.001)
Colour HH SP	0.3775	0.0043	0.8150	0.3950	0.0038	0.0050
N= 2442	(0.103)	(0.004)	(0.126)	(0.206)	(0.006)	(0.006)
White HH SP	0.4875	-0.0005	0.9000	0.4750	0.0017	-0.0023
N= 1865	(0.217)	(0.004)	(0.270)	(0.591)	(0.006)	(0.007)

Parameter estimates from equation – reference – and equation – reference – for all households.

Discussion of Estimates

- For $(a+k)^{\theta}$
    - They match Koch (2017), as they should
    - The esimates are lower than Xu et al. (2003)
    - Meaning: larger equivalence scales
- For $(a+\beta_2 k)^{\beta_1}$
    - Previous Research: $\beta_2 \approx 0.5$, $\beta_1 \approx 0.9$
    - For us: $0.8 \leq \hat{\beta}_2 \leq 1$.
    - But, grid search stopped at 1...
    - For us: $0.37 \leq \hat{\beta}_1 \leq 0.475$
    - Large differences
- Meaning:
    - Children now closer to adult cost than in 1995
    - Household economies more extensive than in 1995
- Scales larger for $(a+\beta_2 k)^{\beta_1}$ than for $(a+k)^{\theta}$

Equivalence Scales for All HH: Models 1 and 2

Table: Equivalence within All Households

		$\Delta = (A+K)^{\theta}$		$\Delta = (A+\beta_1 K)^{\beta_2}$	
Adults	Kids	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)
1	0	1.0000	0.0000	1.0000	0.0000
		(0.000)	(0.000)	(0.000)	(0.000)
1	1	1.3449	0.2963	1.4375	0.3629
		(0.022)	(0.016)	(0.035)	(0.030)
1	2	1.5994	0.4697	1.8750	0.6286
		(0.041)	(0.026)	(0.066)	(0.046)
1	3	1.8088	0.5926	2.3125	0.8383
		(0.059)	(0.033)	(0.096)	(0.056)
2	0	1.3449	0.2963	2.0000	0.6931
		(0.022)	(0.016)	(0.042)	(0.021)
2	1	1.5994	0.4697	2.4375	0.8910
		(0.041)	(0.026)	(0.033)	(0.026)
2	2	1.8088	0.5926	2.8750	1.0561
		(0.059)	(0.033)	(0.046)	(0.035)
2	3	1.9898	0.6880	3.3125	1.1977
		(0.076)	(0.038)	(0.068)	(0.043)

Equivalence scale estimates from equation – reference – and equation – reference – for All households.

Equivalence Scales for Black HH: Models 1 and 2

Table: Equivalence within Black Households

		$\Delta = (A+K)^{\theta}$		$\Delta = (A+\beta_1 K)^{\beta_2}$	
		$\hat{\Delta}$	$\hat{\delta}$	$\hat{\Delta}$	$\hat{\delta}$
Adults	Kids	(s.e.)	(s.e.)	(s.e.)	(s.e.)
1	0	1.0000	0.0000	1.0000	0.0000
		(0.000)	(0.000)	(0.000)	(0.000)
1	1	1.3036	0.2651	1.3925	0.3311
		(0.024)	(0.018)	(0.038)	(0.033)
1	2	1.5223	0.4202	1.7850	0.5794
		(0.044)	(0.029)	(0.073)	(0.052)
1	3	1.6994	0.5303	2.1775	0.7782
		(0.062)	(0.036)	(0.106)	(0.063)
2	0	1.3036	0.2651	2.0000	0.6931
		(0.024)	(0.018)	(0.048)	(0.024)
2	1	1.5223	0.4202	2.3925	0.8723
		(0.044)	(0.029)	(0.037)	(0.028)
2	2	1.6994	0.5303	2.7850	1.0242
		(0.062)	(0.036)	(0.051)	(0.038)
2	3	1.8508	0.6156	3.1775	1.1561
		(0.078)	(0.042)	(0.075)	(0.047)

Equivalence Scales for Coloured HH: Models 1 and 2

Table: Equivalence within Coloured Households

		$\Delta = (A + K)^{\theta}$		$\Delta = (A + \beta_1 K)^{\beta_2}$	
Adults	Kids	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)
1	0	1.0000	0.0000	1.0000	0.0000
		(0.000)	(0.000)	(0.000)	(0.000)
1	1	1.2991	0.2617	1.3117	0.2713
		(0.093)	(0.072)	(0.133)	(0.115)
1	2	1.5140	0.4147	1.6072	0.4745
		(0.172)	(0.113)	(0.250)	(0.180)
1	3	1.6876	0.5233	1.8908	0.6370
		(0.241)	(0.143)	(0.358)	(0.222)
2	0	1.2991	0.2617	1.7593	0.5649
		(0.093)	(0.072)	(0.154)	(0.088)
2	1	1.5140	0.4147	2.0377	0.7118
		(0.172)	(0.113)	(0.181)	(0.112)
2	2	1.6876	0.5233	2.3076	0.8362
		(0.241)	(0.143)	(0.260)	(0.148)
2	3	1.8360	0.6076	2.5706	0.9441
		(0.305)	(0.166)	(0.352)	(0.181)

Equivalence scale estimates from equation – reference – and equation – reference – for Coloured households.

Equivalence Scales for White HH: Models 1 and 2

Table: Equivalence within White Households

		$\Delta = (A+K)^{\theta}$		$\Delta = (A+\beta_1 K)^{\beta_2}$	
		$\hat{\Delta}$	$\hat{\delta}$	$\hat{\Delta}$	$\hat{\delta}$
Adults	Kids	(s.e.)	(s.e.)	(s.e.)	(s.e.)
1	0	1.0000	0.0000	1.0000	0.0000
		(0.000)	(0.000)	(0.000)	(0.000)
1	1	1.4020	0.3379	1.4188	0.3498
		(0.211)	(0.150)	(0.393)	(0.330)
1	2	1.7084	0.5356	1.8240	0.6010
		(0.407)	(0.238)	(0.738)	(0.497)
2	0	1.4020	0.3379	1.8661	0.6238
		(0.211)	(0.150)	(0.350)	(0.187)
2	1	1.7084	0.5356	2.2606	0.8156
		(0.407)	(0.238)	(0.297)	(0.248)
2	2	1.9656	0.6758	2.6475	0.9736
		(0.591)	(0.301)	(0.512)	(0.356)

Equivalence scale estimates from equation – reference – and equation – reference – for White households.

Short Discussion of Race Differentiated Scales

- Ranking: White, Black then Coloured
- Not quite richest to poorest
- Differences not overly large
- But, white estimates noisiest (relatively few observations)
- Still need to complete analysis for each race...
- Policy: Needs more work...
    - VAT exemptions on food?
    - Race differentiated subsidies, taxes and poverty lines?
    - Or, enough to reconsider levels of subsidy, tax and poverty line?
    - Technically, poverty lines not formalized yet in South Africa.
    - Child and spouse income tax exemptions worth consideration.

Outline

Equivalence Scales: All Households Model 3

Table: Child Equivalence within Single-Adult Households

Kids	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)	$\hat{\eta}$ (s.e.)
1	1.6763	0.5166	0.0164
	(0.010)	(0.006)	(0.006)
2	2.7431	1.0091	-0.0179
	(0.010)	(0.004)	(0.009)
3	3.5952	1.2796	-0.0234
	(0.412)	(0.115)	(0.015)

- This set done with subsamples
- But, suggest larger adjustements
- Needs further investigation
- Something interesting with single-adult households...

More Equivalence Scales: All Households Model 3

Table: Child Equivalence within Two-Adult Households

Kids	$\hat{\Delta}$ (s.e.)	$\hat{\delta}$ (s.e.)	$\hat{\eta}$ (s.e.)
1	1.2894	0.2542	-0.0010
	(0.016)	(0.012)	(0.004)
2	1.4773	0.3902	-0.0019
	(0.025)	(0.017)	(0.004)
3	1.9423	0.6639	0.0001
	(0.022)	(0.011)	(0.006)

- This set done with subsamples, too
- Roughly similar estimates

Concluding Thoughts

- Estimated semiparametric equivalence scales assuming base independence
- Race differentiated estimates found
- Race differentiated equivalance scales, too
- Estimates rather different than in 1995:
    - Larger child costs
    - Larger scale economies
    - Overall equivalence similar, but smaller than 1995
- Recently: semiparametric estimates smaller than linear estimates

Final Thoughts

- Recent research suggests a better way forward
  - Browning et al. (2013)
  - GR Dunbar & Pendakur (2014)
  - Chiappori (2016)
- Data requirements, however, exceed what is available
- Although, we do have...
  - Child clothing
  - Female clothing
  - Male clothing
  - Adult clothing
- .. So, maybe?

References I

Anonymous (2016), ‘Inequality, race and equivalences scales in south africa: A research note’, Unpubished Mimeograph, University of the Witwatersrand.

Blackorby, C. & Donaldson, D. (1993), ‘Adult-equivalent scales and the economic implementation of interpersonal comparisons of well-being’, *Social Choice and Welfare* **10**, 335–361.

Blundell, R. & Lewbel, A. (1991), ‘The information content of equivalence scales’, *Journal of Econometrics* **50**, 49–68.

Browning, M., Chiappori, P. & Lewbel, A. (2013), ‘Estimating household economies of scale, adult equivalence scales, and household bargaining power’, *Review of Economic Studies* **80**(4), 1267–1303.

References II

Chiappori, P. (2016), 'Equivalence versus indifference scales', *The Economic Journal* **126**(592), 523–545.

Engel, E. (1857), Die productions- und consumtionsverhaltnisse des konigreichs sachsen, *in* E. Engel, ed., 'Dielebenkostenbelgischerarbeiter-familien', C Heinrich, Dresden.

GR Dunbar, A. L. & Pendakur, K. (2014), 'Identification of random resource shares in collective households with an application to microcredit in Malawi', Unpublished mimeograph, Boston University.

Hayfield, T. & Racine, J. S. (2008), 'Nonparametric econometrics: The np package', *Journal of Statistical Software* **27**(5).
URL: *http://www.jstatsoft.org/v27/i05/*

References III

Koch, S. F. (2017), ‘Does the equivalence scale matter? equivalence and out-of-pocket payments’, Economic Research Southern Africa Working Paper No. 687.

May, J., Budlender, D., Mokate, R., Rogerson, C. & Stavrou, A. (1995), ‘Poverty and inequality in South Africa’, Pretoria: Report prepared for the Office of the Deputy President, Republic of South Africa.

May, J., Carter, M. & Poset, D. (1995), ‘The composition and persistence of poverty in rural South Africa: An entitlements approach’, Land and Agriculture Policy Centre Policy Paper No. 15.

Meth, C. & Dias, R. (2004), ‘Increases in poverty in South Africa, 1999-2002’, *Development Southern Africa* **21**(1), 59–85.

References IV

Pollak, R. A. & Wales, T. J. (1979), ‘Welfare comparisons and equivalence scales’, *American Economic Review* **69**, 216–221.

R Core Team (2016), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Robinson, P. M. (1988), ‘Root-*n*-consistent semiparametric regression’, *Econometrica* **56**(4), 931–954.

Woolard, I. (2002), ‘Income inequality and poverty: methods of estimation and some policy applications for south africa’, PhD Thesis, University of Cape Town.

Woolard, I. & Leibbrandt, M. (1999), ‘Measuring poverty in south africa’, DPRU Working Papers No. 99/33.

References V

Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J. & Murray, C. J. L. (2003), ‘Household catastrophic health expenditure: A multicountry analysis’, *the Lancet* **362**, 111–117.

Yatchew, A., Sun, Y. & Deri, C. (2003), ‘Efficient estimation of semiparametric equivalence scales with evidence from South Africa’, *Journal of Business & Economic Statistics* **21**(2), 247–257.