Online Appendix for "An Estimated Structural Model of Entrepreneurial Behavior"

By John Bailey Jones and Sangeeta Pratap*

This online appendix includes information about data construction, our econometric methodology, and other details that are not included in the main text of the paper.

Appendix A: Data Construction

Sample Selection:

The table below describes the filters used to construct the sample used in the paper from the original data (Dyson School, 2001-2011). Since we are interested in the dynamic behavior of the enterprises, we eliminate farms with only one observation. We also drop farms with missing information on the age of the youngest operator.

	# of Farms	# of Observations
Original Sample	541	2461
Drop farms with one observation	385	2305
Drop farms with missing age	363	2222

Owned Capital:

The sum of the beginning of period market value of the three categories of capital stock owned by a farm: real estate and land, machinery and equipment, and livestock.

Depreciation and Appreciation Rates:

The depreciation rate δ_j for each type of capital stock j is calculated as

$$\delta_j = \frac{1}{T} \frac{\sum_i \text{Depreciation}_{jit}}{\sum_i \text{Market Value of Owned Capital}_{jit}}$$

where i indexes farms and $T = 11$ (2001-2011).

Analogously, the appreciation rate ϖ_j for each type of capital stock j is given by

$$\varpi_j = \frac{1}{T} \frac{\sum_i \text{Appreciation}_{jit}}{\sum_i \text{Market Value of Owned Capital}_{jit}}$$

The actual depreciation rate δ is the weighted average of each δ_j, the weights being the average share of each type of capital stock in owned capital. The appreciation rate ϖ is calculated similarly.

Leased Capital:

* Jones: Federal Reserve Bank of Richmond, 701 E Byrd Street, Richmond VA 23219, john.jones@rich.frb.org. Pratap: Hunter College and Graduate Center, City University of New York, 695 Park Ave, New York NY 10023, sangeeta.pratap@hunter.cuny.edu. The opinions and conclusions are solely those of the authors, and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

The market value of leased capital ($MVLK$) of type j for farm i at time t is calculated as

$$MVLK_{jit} = \frac{\text{Leasing Expenditures}_{jit}}{r + \delta_j - \varpi_j + \pi}$$

where r is the risk-free rate of 4 percent, and π is the average inflation rate through the period. The market value of all leased capital is the sum of the market value of all types of leased capital.

Total Capital :
Owned Capital + Leased Capital

Investment:
Sum of net investment and depreciation expenditures in real estate, livestock and machinery.

Total Output:
Value of all farm receipts.

Total Expenditure:
Expenses on hired labor, feed, lease and repair of machinery and real estate, expenditures on livestock, crop expenditures, insurance, utilities and interest.

Variable Inputs:
Total expenditures less expenditures on interest payments and leasing expenditures on machinery and real estate.

Total Assets:
Beginning of period values of the current assets (cash in bank accounts and accounts receivable), intermediate assets (livestock and machinery) and long term assets (real estate and land).

Total Liabilities:
Beginning of period values of current liabilities (accounts payable and operating debt), intermediate and long-term liabilities.

Dividends:
Net income (total receipts-total expenditures) less retained earnings and equity injections.

Cash:
Total assets less owned capital.

Appendix B: Revenue Distribution

Data for all US firm receipts are available for the Economic Census years (ending in 2 and 7), the last of which was 2012 (United States Census Bureau, 2015). Among single proprietorships, partnerships and S Corporations in the Statistics of US business, average firm receipts in 2012 were \$2.76 million. The farms in our dataset had average receipts of \$2.89 million in 2011, the last year for which we have data.

The distribution of revenues in the U.S. is also comparable with our data, as Figure A1 shows. Since the data are not separated by ownership type, it is not surprising that our sample under-represents firms receiving more than \$15 million annually. It also under-represents very small enterprises, with less than \$100,000 in annual revenue. However, it is quite similar in the middle of the distribution.

Appendix C: Econometric Methodology

We estimate the 12×1 parameter vector $\Omega = (\beta,\ \nu,\ c_0,\ \chi,\ c_1,\ \theta,\ \alpha,\ \gamma,\ n_0,\ \lambda,\ \zeta,\ \psi)$ using a version of Simulated Minimum Distance. To construct our estimation targets, we sort farms along two dimensions, operator age and size (cows per operator). Along each dimension, we

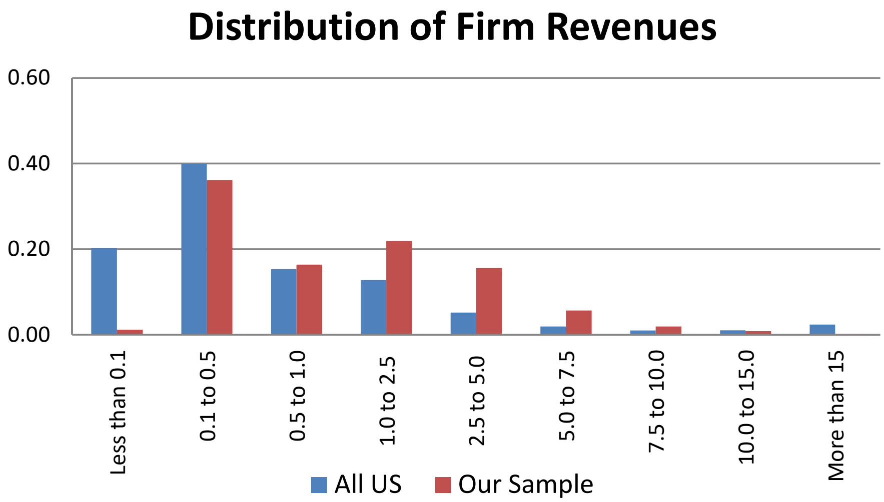

Figure A1. : Distribution of Firm Revenues

divide the sample in half. Then for each of these four age-size cells, for each of the years 2001 to 2011, we match:

1) The median value of capital per operator, k.

2) The median value of the output-to-capital ratio, y/k.

3) The median value of the variable input-to-capital ratio, n/k.

4) The median value of the gross investment-to-capital ratio.

5) The median value of the debt-to-asset ratio, $b/\tilde{a}$

6) The median value of the cash-to-asset ratio, $\ell/\tilde{a}$.

7) The median value of the dividend growth rate, d_t/d_{t-1}.

8) The fraction of farms operating and observed in the DFBS.

Let g_{mt}, $m \in \{1, 2, ..., M\}$, $t \in \{1, 2, ..., T\}$, denote a summary statistic of type m in calendar year t, such as median capital for young, large farms in 2007, calculated from the DFBS. The model-predicted value of g_{mt} is $g^*_{mt}(\Omega)$. We estimate the model by minimizing the weighted squared proportional differences between $\{g^*_{mt}(\Omega)\}$ and $\{g_{mt}\}$.

Suppose we have a sample of I conditionally (on the aggregate shocks) independent farms. Our SMD criterion function is

$$
\begin{aligned}
Q_I(\Omega) &= \sum_{m=1}^{M} \aleph_m \sum_{t=t_1}^{T} \left(\frac{g^*_{mt}(\Omega)}{g_{mt}} - 1 \right)^2 \\
&= [g^*_I(\Omega) - g_I]' \mathbf{W}_I [g^*_I(\Omega) - g_I], \qquad \text{(A1)}
\end{aligned}
$$

where: g_I and g_I^* are $n_{SS} \times 1$ vectors containing the observed and model-predicted summary statistics ($\{g_{mt}\}$ and $\{g_{mt}^*(\Omega)\}$); $\aleph_m$ denotes the weight placed on the mth set of differences; and $\mathbf{W}_I$ is an $n_{SS} \times n_{SS}$ diagonal matrix with typical element $\aleph_m g_{mt}^{-2}$.

Our estimate of the "true" parameter vector Ω_0 is the value of Ω that minimizes the criterion function $Q_I(\Omega)$. Let Ω_I denote this estimate. Under regularity conditions (e.g., Newey and McFadden 1994, Theorem 7.1), the SMD estimator Ω_I is both consistent and asymptotically normally distributed:

$$\sqrt{I}(\Omega_I - \Omega_0) \rightsquigarrow \mathcal{N}(0, \mathbf{V}),$$

with the variance-covariance matrix $\mathbf{V}$ given by

$$\mathbf{V} = (\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{\Sigma}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1},$$

where: $\mathbf{\Sigma}$ is the $n_{SS} \times n_{SS}$ asymptotic variance-covariance matrix of $\sqrt{I}(g_I^*(\Omega_0) - g_I)$; and

$$\text{(A2)} \qquad \mathbf{D} = \left.\frac{\partial g^*(\Omega)}{\partial \Omega'}\right|_{\Omega=\Omega_0}$$

is the $n_{SS} \times 12$ Jacobian matrix of the population summary statistics; and $\mathbf{W} = \text{plim}_{I\to\infty}\{\widehat{\mathbf{W}}_I\}$.

We estimate $\mathbf{W}$ with its sample analog $\mathbf{W}_I$. We find $\mathbf{D}$ by taking the numerical derivative of $g_I^*(\Omega)$ at Ω_I. Because numerical gradients are sensitive to step sizes, we follow the approach used by Bloom (2009). For each parameter, we calculate the slope using steps equal to 5%, 2.5%, 1% and -1% of the parameter in question and keep the median value.

Unfortunately, there is no analytical expression for the limiting variance $\mathbf{\Sigma}$. We instead find $\mathbf{\Sigma}$ via a bootstrap procedure. Because the model simulations use the estimated values of the aggregate shocks $\{\Delta_t\}$ and the farm-specific effects $\{\mu_i\}$, variation in these estimates will affect the model predictions ($g_I^*(\Omega_0)$) as well as the data (g_I). In linear panel regression, the standard practice is to purge the data of fixed and time effects; this is common practice in structural analyses as well (e.g. Strebulaev and Whited, 2012). In contrast, we are comfortable giving these effects a strict structural interpretation, because of the quality and detail of the DFBS. Including these effects, however, requires us to account for them in our bootstrap procedure.

To account for idiosyncratic variation, we create $S = 400$ artificial samples of size I, each sample consisting of I bootstrap draws from the DFBS. Each draw contains the entire history of the selected farm. In this respect, we follow Kapetanios (2008), who argues that this is a good way to capture temporal dependence in panel data bootstraps. For each bootstrapped sample $s = 1, 2, \ldots, S$, we compile the vector of summary statistics for the data sample, g_{Is}, and the vector of summary statistics predicted by the model, $g_{Is}^*(\Omega_I)$, and record the difference $[g_{Is}^*(\Omega_I) - g_{Is}]$. While every bootstrapped value of $g_{Is}^*(\Omega_I)$ is calculated using the parameter vector Ω_I, variation in the bootstrapped distribution of farm-specific effects, $\{\mu_i\}$, and the simulated idiosyncratic shocks, $\{\varepsilon_{it}\}$, generates variation in $g_{Is}^*(\Omega_I)$.

To account for variation due to aggregate shocks. we assume that the time-t observation of the type-m summary statistic can be written as

$$g_{mt} = a_{0m} + a_{1m}t + a_{2m}t^2 + \epsilon_{mt},$$

where ϵ_{mt} is independent across t but not necessarily across m. We include a quadratic time trend because the moments are clearly nonstationary. This gives us $T - 1$ observations of the $n_{SS} \times 1$ residual vector ϵ_t. (Because some moments are not observed in the first period, this residual is discarded.) We then add to each realization of g_{Is} a set of T bootstrapped draws of ϵ_t, drawing complete $n_{SS} \times 1$ vectors to preserve the correlation structure of the summary statistics. At this point each value of g_{Is} reflects aggregate as well as idiosyncratic variation. Finding the variance of g_{Is} (inclusive of aggregate shocks) across the S subsamples yields $\mathbf{\Sigma}_S$, our estimate of $\mathbf{\Sigma}/I$ (not $\mathbf{\Sigma}$).[1]

While our bootstrap estimate of $\mathbf{\Sigma}$ captures the effect of aggregate shocks on the sampling variation of the data statistic g_I, it does not capture the effect of aggregate shocks on the sampling variation of the model prediction $g_I^*(\Omega_0)$. We know of no good way to do so. It is possible that accounting for this effect would *reduce* the estimated variance of $(g_I^*(\Omega_I) - g_I)$, as aggregate shock-related variation in $g_I^*(\Omega_I)$ might offset aggregate shock-related variation in g_I. If this is the case, then our bootstrap estimator will overstate $\mathbf{\Sigma}$ and the parameter variance $\mathbf{V}$.

	Medians		Means		Std. Deviation†		Autocorrelations†	
	Data	Model	Data	Model	Data	Model	Data	Model
Capital	586	512	845	781	959	876	0.98	0.99
Y/K	0.38	0.38	0.40	0.43	0.14	0.18	0.86	0.90
N/K	0.28	0.25	0.30	0.30	0.12	0.12	0.85	0.92
I/K	0.04	0.03	0.07	0.08	0.26	0.17	0.08	0.27
Debt/Assets	0.49	0.51	0.51	0.45	0.21	0.25	0.92	0.86
Cash/Assets	0.10	0.10	0.11	0.11	0.05	0.05	0.84	0.57
Dividend growth	0.86	1.00	0.73	1.02	4.79	0.08	0.02	0.16

*Standard deviations and autocorrelations use deviations from annual means.

Table A1—: Data and Baseline Model Moments (weighted by herd size)

Appendix D: Goodness of Fit

Tables A1 and A2 compare data moments from the DFBS with moments for data generated by the baseline model. The criterion function targets yearly medians directly, but the model means and standard deviations are very similar to those in the data. A notable exception is the variability of the growth rate of dividends, which is much larger in the data than in the model. This is partly due to the presence of outliers: dropping the top and bottom 1 percent of the sample gives us a standard deviation of 1.97, while keeping the mean virtually unchanged. However, as discussed in the main text (see Section IV.B), our CRRA utility function cannot generate the low degree of intertemporal substitutability implied by the low observed dividend growth rate and also generate the low degree of risk aversion implied by the high observed dividend volatility. An alternative specification where we used Epstein-Zin style (1989) preferences to unlink intertemporal substitutability and risk aversion did not materially improve our results.

The model also does a reasonable job in matching the autocorrelations found in the data. The cross correlations between the real and financial variables are mostly captured by the model as well.

Appendix E: Modelling the 2014 Farm Bill

The 2014 Farm Bill program replaced the previous dairy support program, which guaranteed a minimum price for milk, with a guarantee of the milk margin, the difference between the price of

[1] Because $g_{mt}^*()$ is found via simulation rather than analytically, the variance Σ must account for simulation error. In most cases the adjustment involves a multiplicative adjustment (Pakes and Pollard, 1989; Duffie and Singleton, 1993; Gouriéroux and Monfort, 1996). Because each iteration of our bootstrap employs new random numbers, no such adjustment is needed here.

	Capital	Y/K	N/K	I/K	Debt/ Assets	Cash/ Assets	Dividend Growth
Data							
Capital	1.00	0.19	0.24	-0.01	0.07	0.18	0.00
Y/K		1.00	0.94	0.21	0.19	0.57	0.05
N/K			1.00	0.21	0.17	0.55	0.02
I/K				1.00	0.13	0.06	0.01
Debt/Assets					1.00	-0.01	-0.01
Cash/Assets						1.00	0.03
Dividend Growth							1.00
Model							
Capital	1.00	0.02	0.07	-0.07	0.18	0.08	-0.06
Y/K		1.00	0.96	0.60	0.44	0.49	0.47
N/K			1.00	0.49	0.40	0.55	0.37
I/K				1.00	0.23	0.25	0.53
Debt/Assets					1.00	0.02	0.15
Cash/Assets						1.00	0.29
Dividend Growth							1.00

Table A2—: Data and Baseline Model Cross Correlations (weighted by herd size)

Notes: Correlations calculated using deviations from annual means.

milk and a weighted average of the prices of corn, soybean and alfalfa. As described by Schnepf (2014), the program provides a baseline margin support of \$4.00 per cwt. for all participants. Higher support levels (in \$0.50 increments) can be acquired in exchange for a premium. For the first 4 million lbs of output, the premium ranges from \$0.01 per cwt. for a margin of \$4.50 to \$0.475 for a \$8.00 margin. For additional output the premium ranges from \$0.02 to \$1.36 per cwt.

As discussed in the main text, we model margin floors by eliminating the left-hand tail of the aggregate shock distribution. Schnepf (2014) notes that the premium structure encourages farmers to choose a margin support level of \$6.50 per cwt. In the implementation of the margin support program, the milk margin is computed and compared to the margin floor every two months. At this frequency, during our sample period of 2001 to 2011 nominal milk margins fell beneath \$6.50 about 13.6 percent of the time. Real milk margins, measured in September 2014 dollars, fell beneath the floor about 7.6 percent of the time. At the annual frequency used in our model, both nominal and real milk margins fell below the \$6.50 floor once, in 2009, a rate of $1/11 = 9.09$ percent. We choose this annual figure as our truncation level.

Following standard practice, when solving the model numerically we replace the continuous processes for the productivity shocks with discrete approximations. We use the approach developed by Tauchen (1986) for discretizing Markov chains, simplified here to i.i.d. processes. Under Tauchen's approach one divides the support of the underlying continuous process into a finite set of intervals. The values for the discrete approximation come from the interiors of the intervals (demarcated by the midpoints, except at the upper and lower tails), and the transition probabilities are based on the conditional probabilities of each interval. To impose the cutoff, we construct the discretization for the aggregate shock so the bottom two states/intervals have a combined probability of 9.09 percent. We then set the truncated value for these states to equal the 9.09th percentile of the underlying continous distribution.

When finding the decision rules, we combine the aggregate and idiosyncratic shocks into a single transitory shock. In the baseline specification, which we use when estimating the model, we approximate the sum of the two-shock processes with an eight-state discretization. To model the elimination of certain shocks, we simply change the standard deviation of this combined process. However, to capture the Farm Bill, we need to alter the distribution of the aggregate shock in isolation. We thus model the aggregate shock with its own eight-state discretization. To get the joint shock, we approximate the idiosyncratic shock with a four-state distribution, and convolute the aggregate and idiosyncratic shocks into a 32-state distribution. We construct the four-state discretization so that the standard deviation of the convoluted 32-state process is the same as the standard deviation of the eight-state shock used in the baseline model. Switching between the two discrete approximations affects the model's results only modestly (compare the first lines of Tables 5 and 7 in the main text). Under either approximation, the shocks used in the simulations are a combination of idiosyncratic shocks from a random number generator and the aggregate shocks estimated from data. To simulate the Farm Bill, we simply truncate the aggregate shock for 2009 to the 9.09th percentile.

For large volumes, the premium for a margin support of \$6.50 is \$0.29 per cwt. This equals about 1.58 percent of the average real milk price over the 2001-2011 interval (\$18.34). We impose the premium by incrementing the logged productivity process by $\ln(0.9842)$, that is, by reducing the productivity level by 1.58 percent.

Appendix F: Weighted Estimation

In our comparative statics exercises we apply herd size weights to the simulation output, resulting in averages that are representative for New York State. Under our of assumption of exogenous stratification, however, applying these weights while estimating the model would be inefficient (Wooldridge, 1999), and our estimation targets are thus based on the unweighted data. Here we discuss the estimates from a specification where we apply herd size weights. We split the sample as in our baseline approach but: (1) the summary statistics within each size-year cell are calculated using size weights; (2) the weight placed on each size-year cell in the GMM criterion function depends on the size weights.

Column (1) of Table A3 shows the results. Using the weights results in an estimate of ν consistent with a high degree of risk aversion and an estimate of ψ consistent with a weak borrowing constraint. We find these estimates less plausible than our baseline estimates, for a number of reasons. First, our weighting procedure is a crude one based on a small set of farm size bins. Second, the model with the weighted parameter estimates is unable to explain the investment behavior of young, high-productivity farms; this can be seen in the upper left corner of Figure A2. Third, the estimates are not robust. If we match the weighted mean of capital holdings, as opposed to the median, the weighted estimation procedure delivers parameters akin to the baseline estimates. Column (2) of Table A3 shows these estimates. Finally, the weighted estimates imply that farmers would value milk price insurance very highly, and yet we do not observe them buying such products, even though futures contracts for milk, fuel and grain are all traded in commodities markets. There may be features of these contracts that discourage their use, but it bears noting that participation in the milk margin support program has been limited (Johnson, 2017).

Parameter Description		Baseline Moments (1)	Mean Capital (2)
Discount factor	β	1.008	0.973
		(0.023)	(0.018)
Risk aversion	ν	10.252	4.355
		(0.294)	(0.249)
Consumption utility shifter (in \$000s)	c_0	4.214	3.673
		(0.228)	(0.300)
Retirement utility shifter (in \$000s)	c_1	18.03	16.17
		(0.493)	(0.677)
Retirement utility intensity	θ	117.0	107.6
		(3.15)	(3.40)
Nonpecuniary value of farming (consumption decrement in \$000s)	χ_C	6.588	6.296
		(0.268)	(0.306)
Returns to management: stanchions	α	0.116	0.126
		(0.002)	(0.002)
Returns to management: parlors	α	0.114	0.113
		(0.002)	(0.002)
Returns to capital: stanchions	γ	0.183	0.176
		(0.002)	(0.004)
Returns to capital: parlors	γ	0.123	0.121
		(0.002)	(0.003)
Strength of collateral constraint	ψ	0.803	1.031
		(0.016)	(0.028)
Degree of liquidity constraint	ζ	3.244	2.648
		(0.078)	(0.067)

Table A3—: Parameters Estimated Using Herd Size Weights

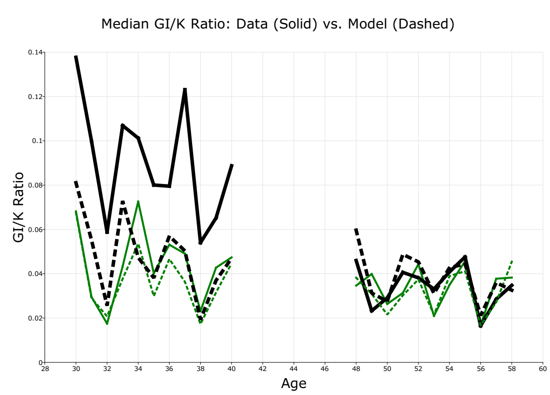

Figure A2. : Observed and Model-predicted Investment Rates under Herd Size Weighting

Note: Solid lines refer to weighted DFBS data, dashed lines to weighted model simulations. Thicker lines refer to farms with large herds, thinner lines refer to farms with small herds.

REFERENCES

Bloom, Nicholas, 2009. "The Impact of Uncertainty Shocks," *Econometrica* 77(3): 623-685.

Duffie, Darrell and Kenneth J. Singleton, 1993. "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica* 61(4): 929-952.

Dyson School of Applied Economics and Management, 2001-2011. "Dairy Farm Business Survey," Cornell University. `http://dfbs.dyson.cornell.edu.`

Epstein, Larry G., and Stanley E. Zin, 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57(4): 937-969.

Gouriéroux, Christian, and Alain Monfort, 1996. *Simulation-Based Econometric Methods*, Oxford: Oxford Univ. Press.

Johnson, Renée, 2017. "Previewing a 2018 Farm Bill", CRS Report March 2017.

Kapetanios, George, 2008. "A Bootstrap Procedure for Panel Data Sets with Many Cross-sectional Units," *Econometrics Journal* 11(2): 377-395.

Newey, Whitney, and Daniel McFadden, 1994. "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, edited by Robert Engle and Daniel McFadden Vol. 4: 2111-2245, Amsterdam: Elsevier.

Pakes, Ariel and David Pollard, 1989. "Simulation and the Asymptotics of Optimization Estimators," *Econometrica* 57(5): 1027-1057.

Schnepf, R., 2014. "Dairy Provisions in the 2014 Farm Bill (P.L. 113-79)," CRS Report, September.

Strebulaev, Ilya A., and Toni M. Whited, 2012. "Dynamic Models and Structural Estimation in Corporate Finance," *Foundations and Trends in Finance* 6(1-2): 1-163.

Tauchen, George, 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters* 20: 177–181.

United States Census Bureau, 2015. "2012 SUSB Annual Data Tables by Establishment Industry," available at `https://www.census.gov/data/tables/2012/econ/susb/2012-susb-annual.htm`.

Wooldridge, Jeffrey M, 1999, "Asymptotic Properties of Weighted M estimators for Variable Probability Samples," *Econometrica* 67(6): 1385-1406.