Small area population projections using stochastic simulation

Stuart H. Sweeney
Department of Geography
University of California, Santa Barbara

December 2002

Applied Demography Seminar, University of Wisconsin at Madison

Overview

- Motivation
- Uncertainty in population forecasting
- Stochastic simulation for the HU method
- Application: SCAG (dynESDA demo)
- Ongoing research

Motivation

- Research contract:
  - Southern California Association of Governments (SCAG)
  - 2030 population/household projections for transportation plan

- Objectives:
  1. Projection of observed trends to 2030 (no policy/development scenarios)
  2. Output: *Geography* - city, unincorporated subregion/RSA
     Variables - total, residential, and GQ population
     - households by type of housing unit (SF, MF, MB)
  3. Develop new methodologies for future forecasting exercises.

Motivation (cont.)

- Constraints:

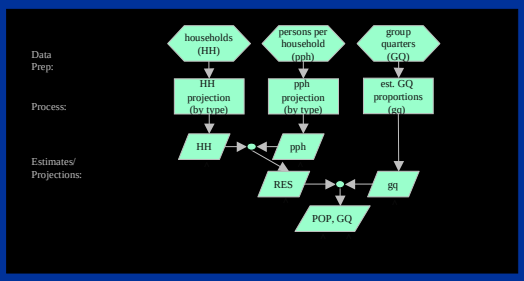
  Methodology:
  - Overall framework must be the housing unit method
  - County-level forecasts are taken as given and inviolable

  Political and administrative context:
  - All output trends must be non-negative
  - Project households directly, not housing units
  - Final projections should accord with expert judgment

  Data:
  - Can only use data available in the SCAG data library

Motivation (cont.)

- Housing unit method:
  - literature tends to focus on intercensal estimates, not projections
  - methodological stagnancy (Voss)
  - output is usually in the form of a single best projection
    - *"Projection, forecast, and plan" or "Scenarios, bluff, and fold"*
  - no known attempts to incorporate uncertainty into HU method



Uncertainty in population forecasting

- Why?

"Current official population forecasts differ little from those that Whelpton made 50 years ago either in the cohort-component methodology used or in the arguments used to motivate the assumptions. However, Whelpton produced some of the most erroneous forecasts of this century. This suggests that current forecasters should ensure that they give users an assessment of the uncertainty of their forecasts."

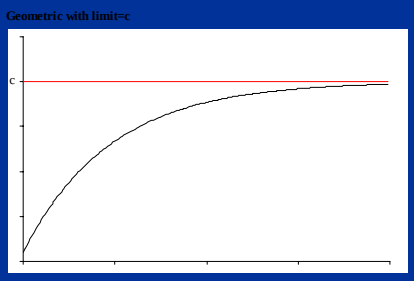
"...current official demographic forecasts are typically judgmental and in most countries there is no systematic effort to provide the forecast user with a probabilistic assessment of the uncertainty of the forecasts."

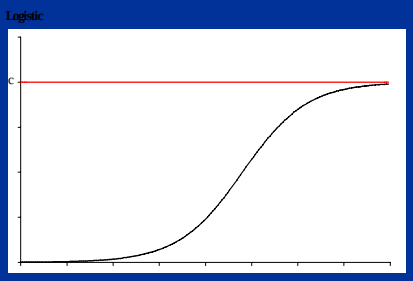
- Juha Alho (1997)

Stochastic simulation for the HU method

- Housing units by type
  - – single-family, multi-family, mobile home, other.

- Constrained growth path





Stochastic simulation for the HU method (cont.)

- Generalized constraint vector, *C*.

$$C \propto f(W,L,I,G)$$

$$C \propto \tilde{W}^{\alpha_1} + \tilde{L}^{\alpha_2} + \tilde{I}^{\alpha_3} + \tilde{G}^{\alpha_4}$$

W=water, *L*=land, *I*=institutional context, *G*=past growth

α= subjective or estimated weights

- Probability distributions

$$p_i = c_i / c_+$$

$$C^{(z)} \sim Dr(P)$$

$$P^{(z)} \sim beta(w,v)$$

Stochastic simulation for the HU method (cont.)

- Simple case with no policy variables, $C \propto f(G)$:
- Algorithm

Step 1: Define G=$\{g_5, g_{10}, g_{15}\}$

Step 2: $P^{(z)}$= [min(G)+rng(G) beta(v,w)]

Step 3: $\hat{C}^{(z)} = Dr(P^{(z)})$

Step 4: $\hat{Y}^{(z)} = pred(Y, C^{(z)})$

Step 5: repeat (2)-(4) *Z* times

- Each complete set of city projections is consistent with county control total... almost.

Stochastic simulation for the HU method (cont.)

Other issues:

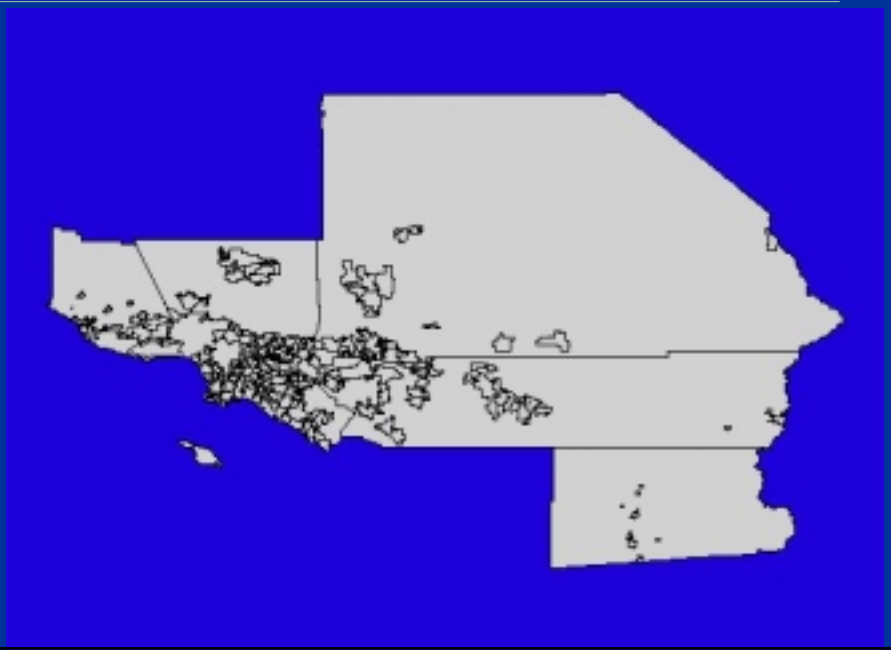
A. For housing type, *i*, iteration (z) yields a complete spatial-temporal panel.

Step 6: Select single prediction, *min* MAD from median.

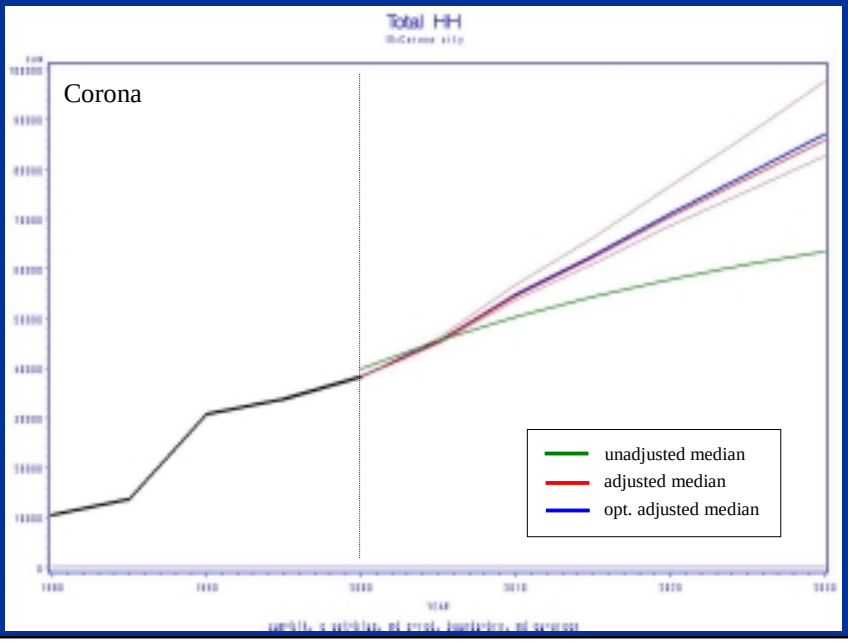
B. Estimation for housing types is independent of other types

Step 7: IPF algorithm

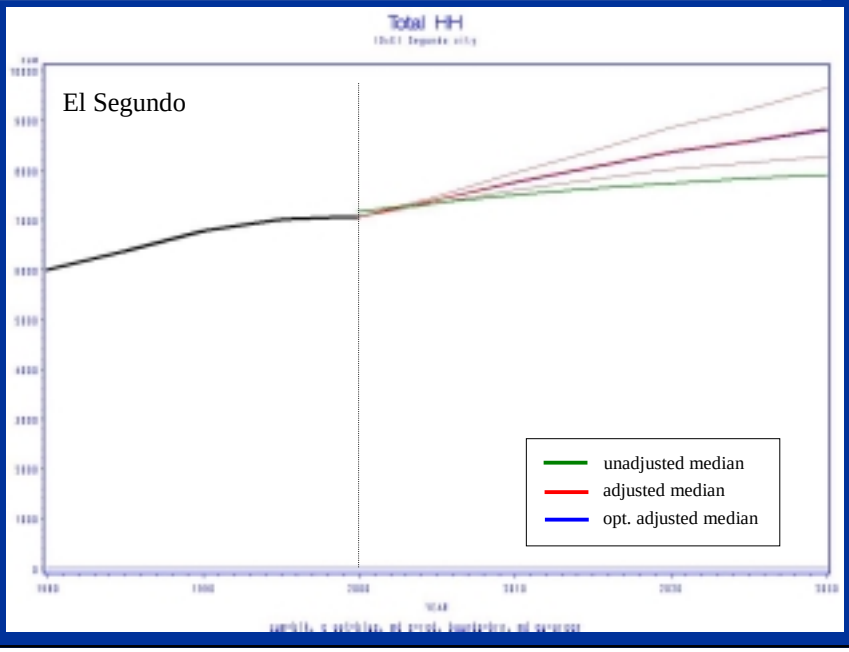
Stochastic HU method: Sample Results



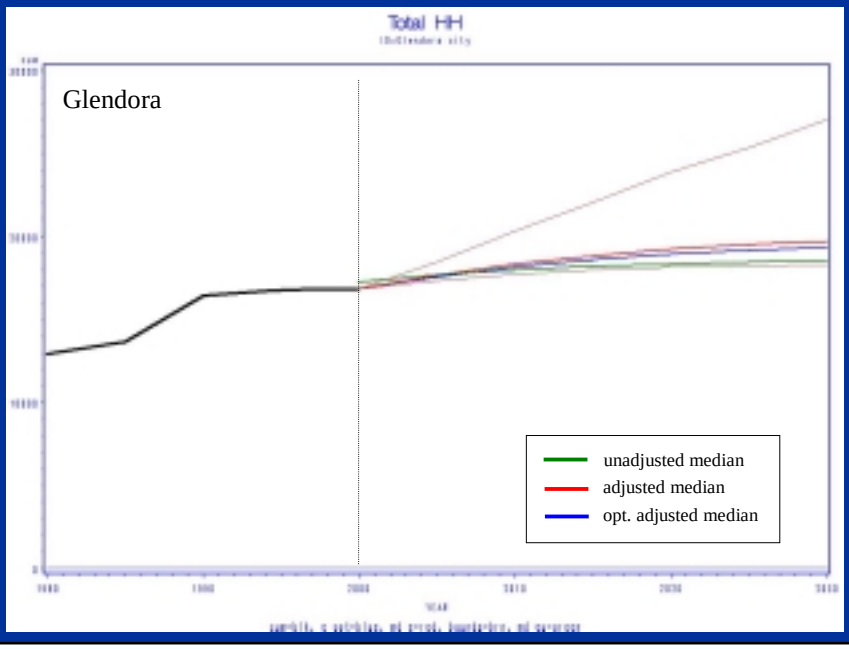
Stochastic HU method: Sample Results



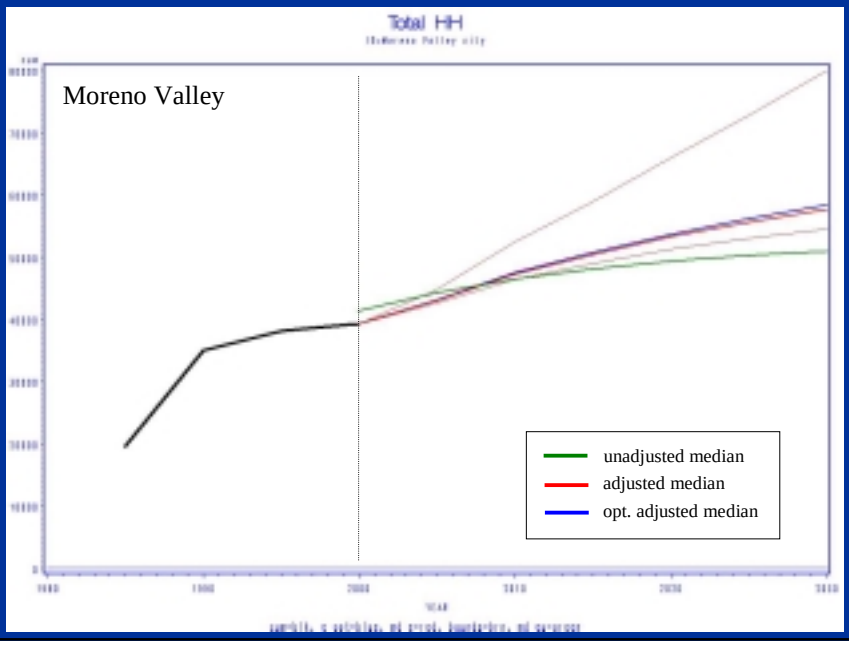
Stochastic HU method: Sample Results



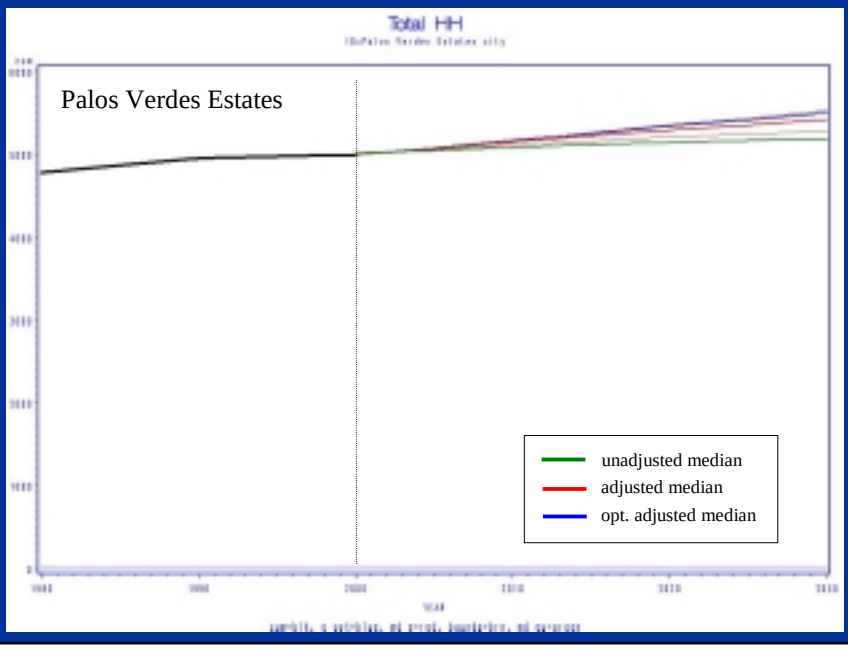
Stochastic HU method: Sample Results



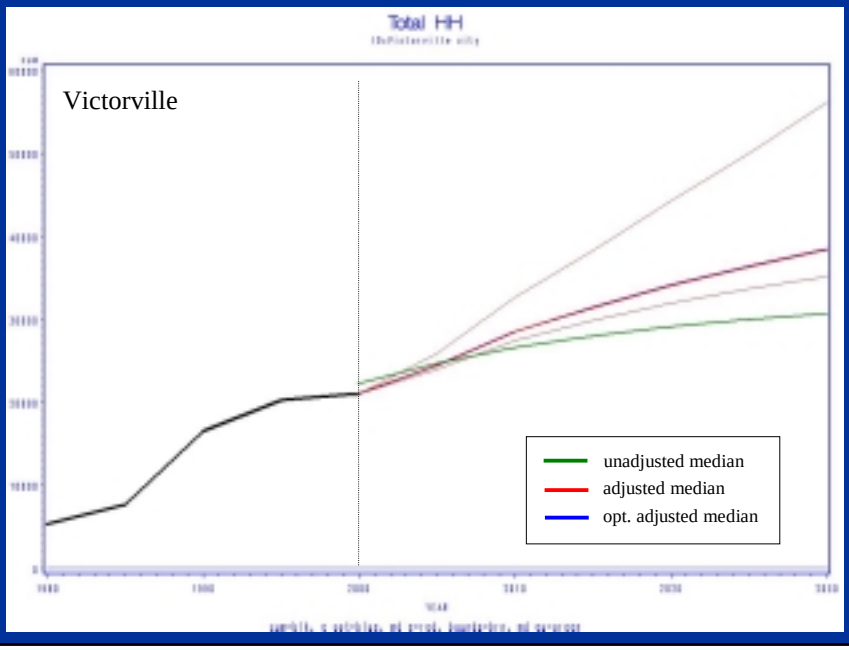
Stochastic HU method: Sample Results



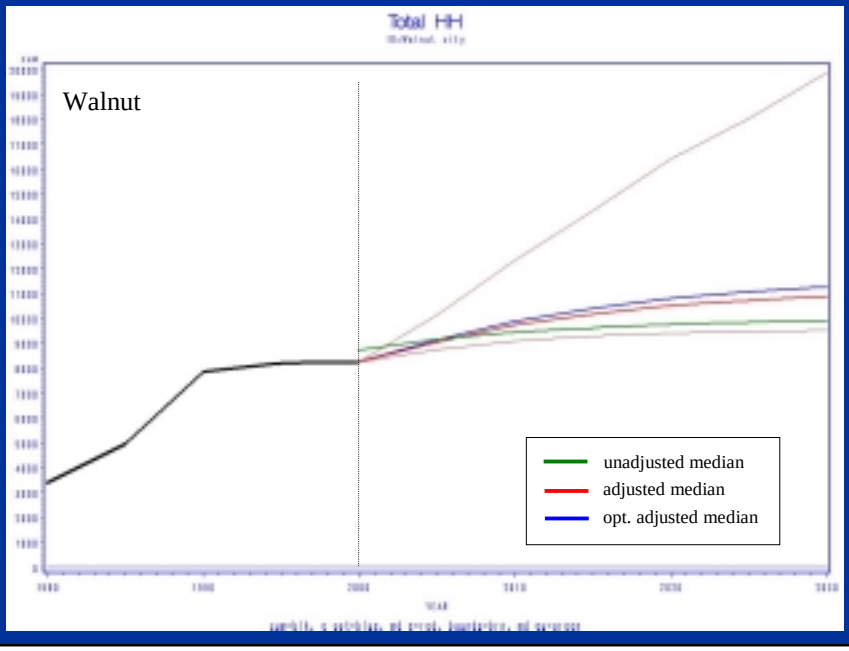
Stochastic HU method: Sample Results



Stochastic HU method: Sample Results



Stochastic HU method: Sample Results



Output evaluation: dynESDA

- Measures
  - *Growth rates: r_O and r_P*
  - *Households in 2000 (h0)*
  - *Uncertainty: (U99-L99)/.5*(U99+L99)*
  - *Deviation in medians: (UM- AM)/.5*(UM+AM)*
- DynESDA – chloropleth, scatterplot, boxplots, local & global I

Ongoing research

- Stochastic specification
- Decompose error into model versus constraint
- Alternative constraints
- Hierarchical Bayes small area estimators with unconstrained as auxilliary information
- Communication of uncertainty to policymakers and planners