The Insights and Illusions of Consumption Measurement:

Evidence from a Large Scale Experiment

Erich Battistin
University of Maryland, IRVAPP, CEPR and IZA

Michele De Nadai
University of New South Wales

Nandini Krishnan
World Bank

IARIW World Bank Conference
November 7-8, 2019

What We Do

- Revisit if **diaries** are the most appropriate benchmark at a time when developing countries are increasingly moving to **recall**.
- Leverage a unique large-scale **pilot** in Iraq to inform the national statistics agency on the transition from diary to recall.
- Identify measurement error distributions in consumption from both diary and recall measurements:
  - Allow for **errors in both measurements**.
  - Errors from alternative collection modes can be correlated and correlated with true consumption.
- Our assumptions hinge on the same survey design employed in many national household surveys, like in Canada and the United States.

Four Take-Home Messages

1 - The presumption that diaries are the most appropriate instrument to measure consumption is an **illusion**.
   - The modal entry understates true consumption by about 20%, with a thick lower tail: $13\% - 16\%$ reporting less than half.
   - Severe over-reporting (share reporting at least twice) is 5%.

2 - Recall errors are far from being classical in form, over-reporting being more likely than under-reporting.
   - However, the modal entry is about right.
   - Severe under-reporting (share reporting at most half) is in the $1\% - 3\%$ range.

3 - **Recall data yield a better classification of the poor**, and poverty statistics are severely overstated using diaries.

4 - What is the value of one additional dollar spent on a diary?
   - An approximately **equal mix** of diary and recall interviews yields empirical distributions closest to the true distribution.
   - A recall survey is the best option when little information on true household consumption is available.

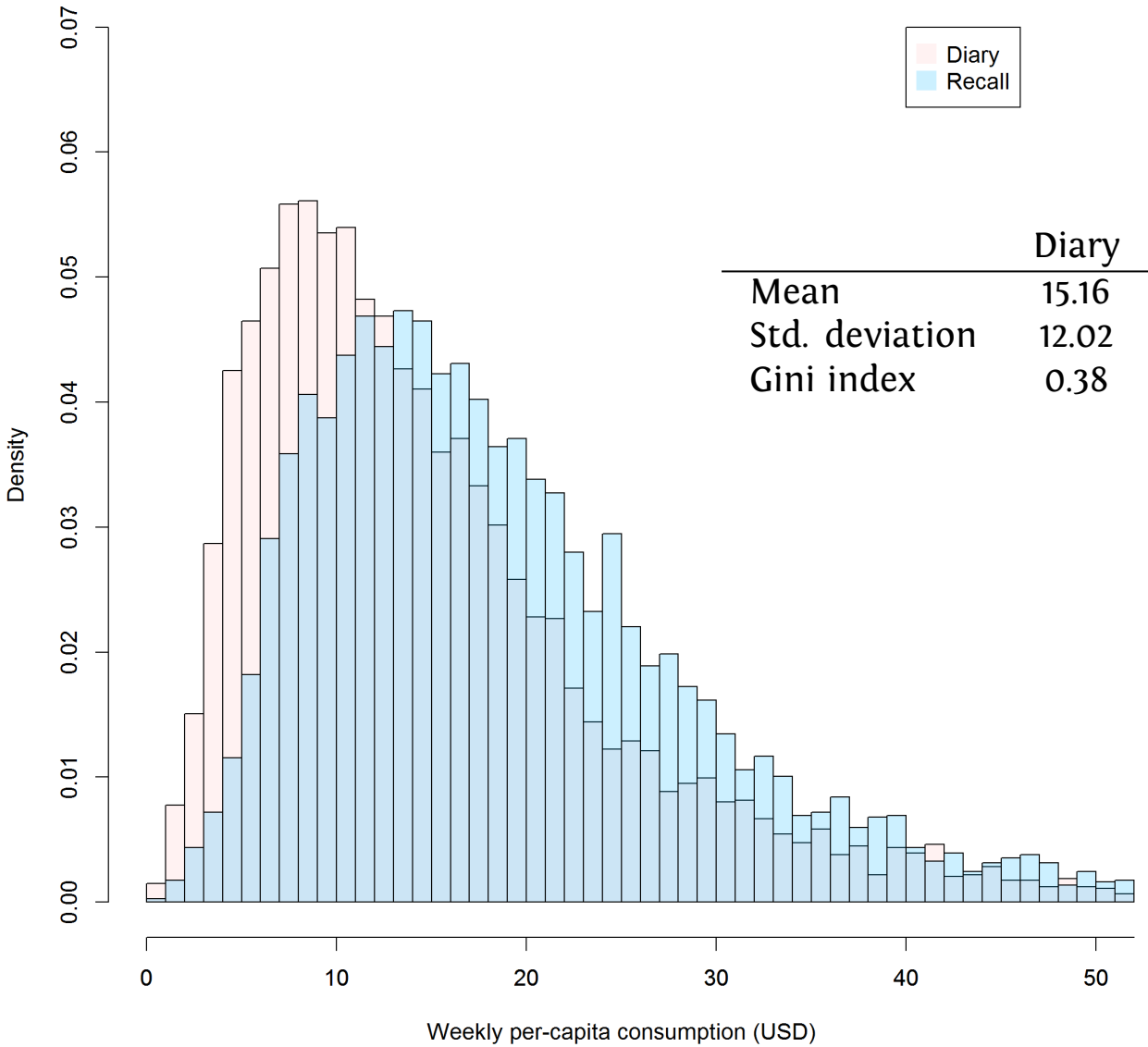
The Iraq Household and Socio Economic Survey

- We work with the 2012 Living Standards Measurement Survey (LSMS), covering about 25 000 households.
  - One-year long survey.
  - Monthly interviews in randomly selected EAs, with 9 households each.
- The baseline mode of collection is a **one-week household diary** filled out with the external assistance of enumerators (five visits).
- Diaries capture **valued acquisition**. Information on source (market, grants, donations, and gifts), amount, quantity and units purchased.
- In addition, households 3, 6 and 9 in each EA are administered a **recall module** on food **consumption** prior to the recording of diaries.
- By design:
  - 25 000 households fill out a diary.
  - One-third of these households, randomly selected, are mandated to an additional recall module.

The Recall Module

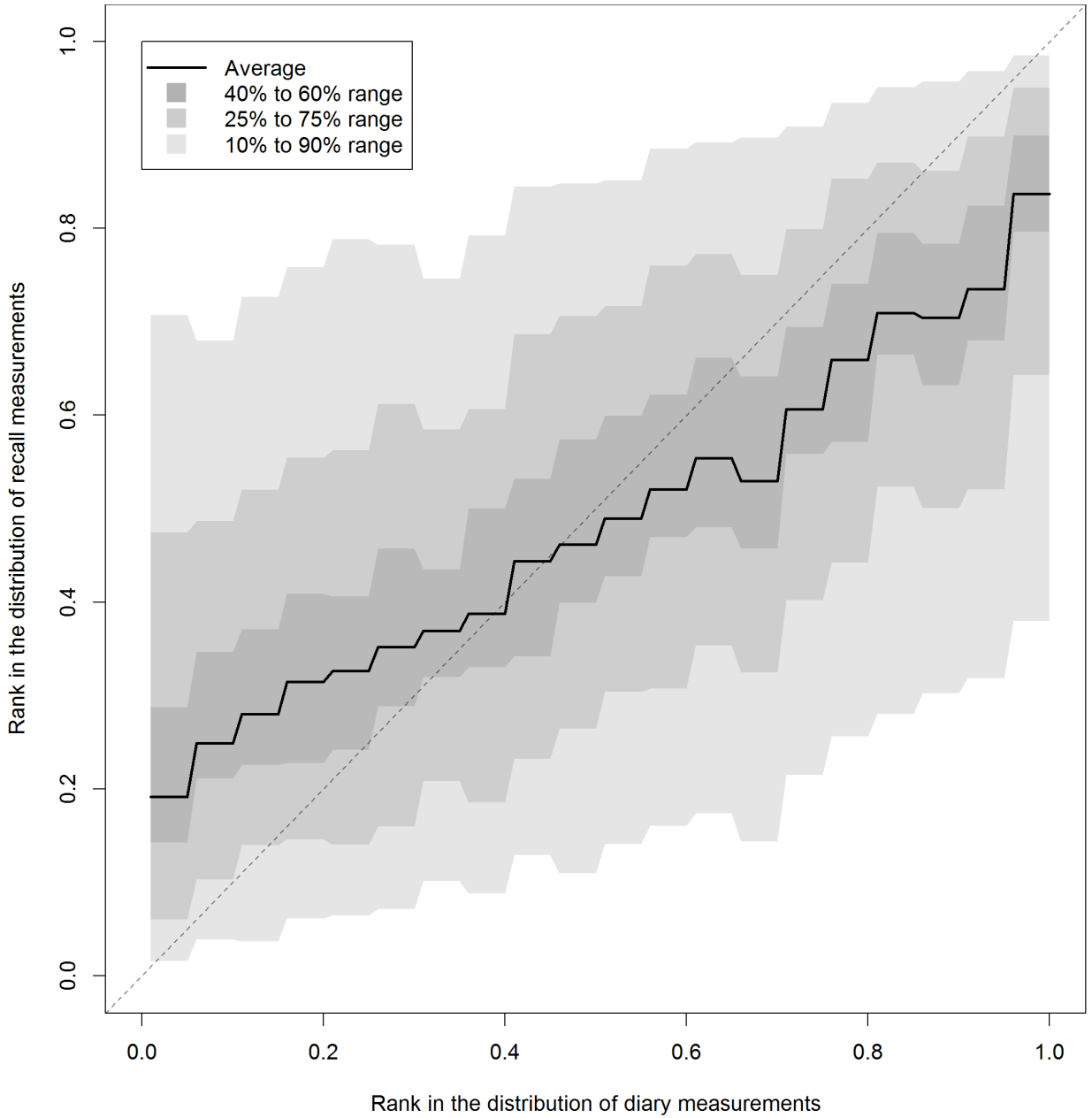
- Specifically designed to inform the national statistics agency on the transition from diary to recall, planned for 2020.
- Refers to the **week before** the first visit.
- Uses a list of **20 groups** obtained from disaggregated categories.
- The list was selected based on an assessment of their importance in household food budget shares and on how commonly they were reported, based on IHSES diaries for 2007.
- In the analysis:
  - Further **harmonization** of household measurements.
  - Randomization of the recall module was successful as well as the **randomization** of households to interviews in **different survey months**.

Data: Recall vs Diary



	Diary	Recall
Mean	15.16	18.82
Std. deviation	12.02	11.53
Gini index	0.38	0.31

Data: Recall vs Diary Rank



A Matter of Frequency

- Simple simulations can give a sense for the difference between **spending** and **consumption** in the diary survey.
  - Y^* is weekly household consumption.
  - N^* is the **typical** number of purchases over one week.
  - N is the **observed** number of purchases in the diary (0 to 3 in our data).
  - To fix ideas, assume purchases of equal amounts.
- Diary spending (Y^d) is:

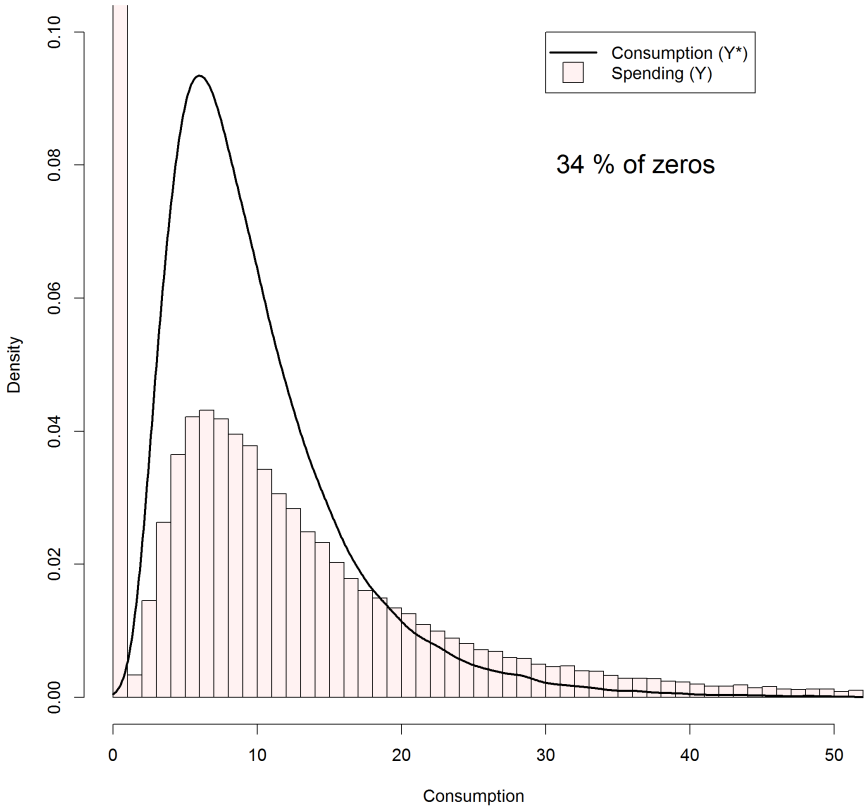
$$Y^d = \frac{Y^*}{N^*} N$$

- Running example: household consuming $\$40$ worth of chicken per month, purchasing it 8 times per month:
  - $Y^* = 10$, $N^* = 2$.
  - Each purchase is $\$5 = 10/2 = 40/8$.
- A reasonable assumption is $E[N|N^*, Y^*] = N^*$. Note that this implies

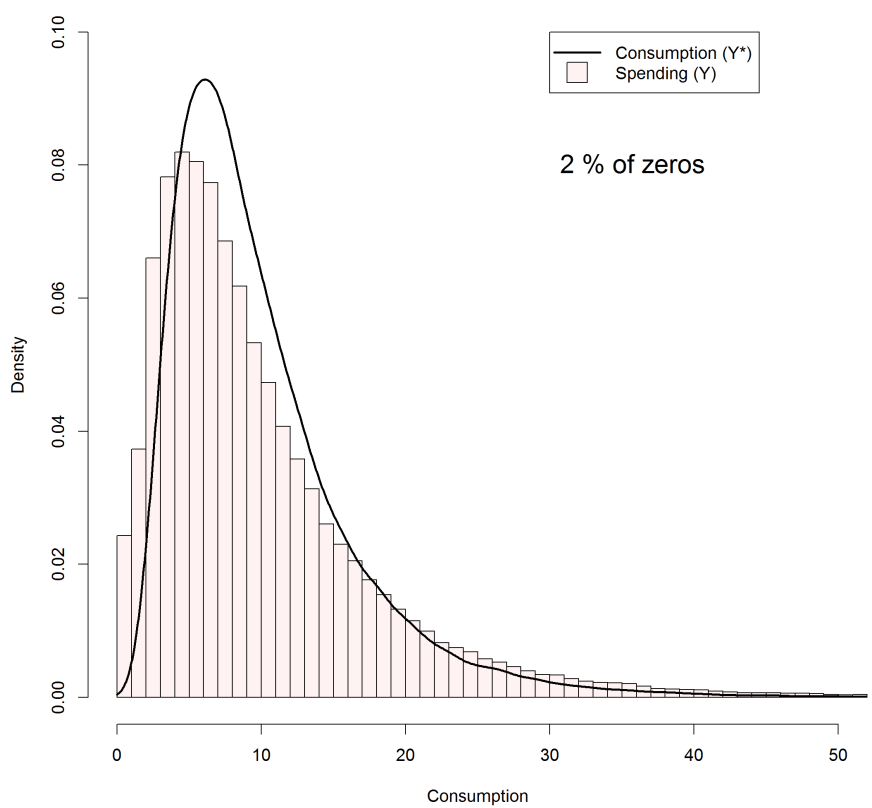
$$E[Y^d] = E[Y^*]$$

Simulation: Consumption vs. Spending - One Item

$N^* = 1$



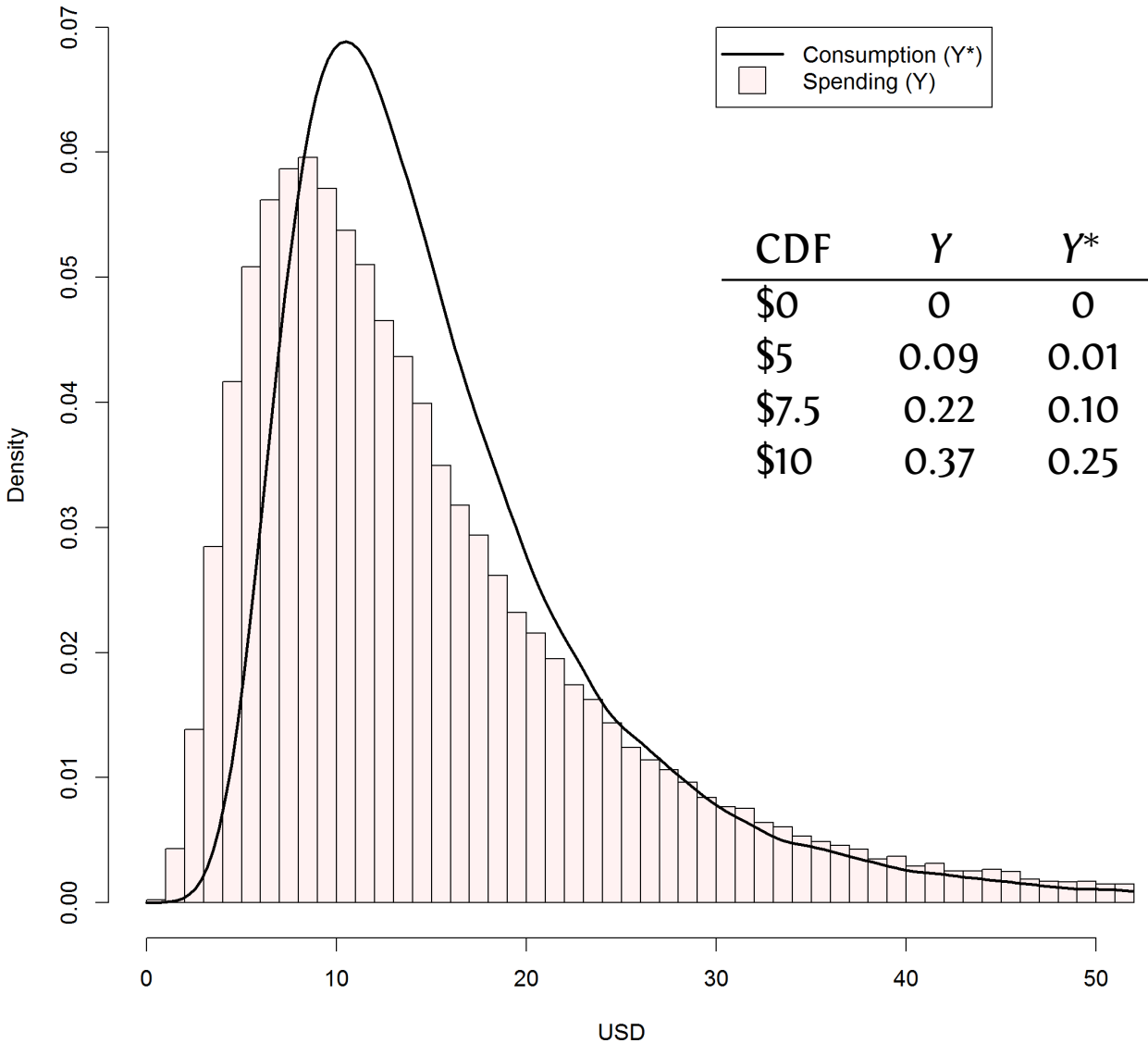
$N^* = 3$



- 15 out of 20 items considered in the analysis are bought **less than once per week**.

Simulation: "Consumption" vs. "Spending" — Sum Over Items

Aggregation over Goods Mitigates Problem

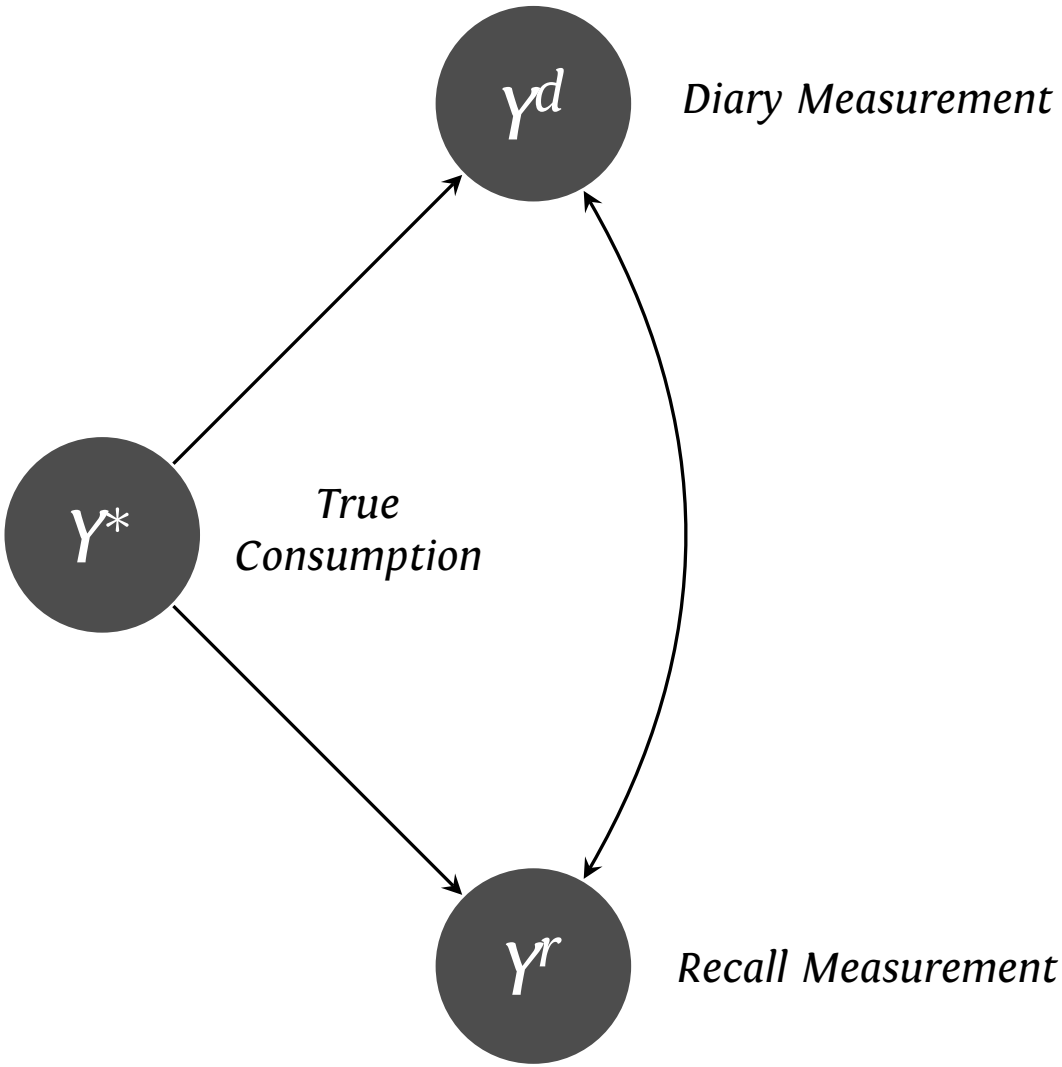


CDF	Y	Y^*	Diff.
\$0	0	0	0
\$5	0.09	0.01	0.08
\$7.5	0.22	0.10	0.12
\$10	0.37	0.25	0.12

Experiments Meet Repeated Measurements

- The Iraq setting allows **identification of distributions** of both latent consumption Y^* and measurement errors in diaries and recall.
  - No need to assume that diary measurements are error-free.
  - Errors in diary and recall measurements can depend on Y^*.
  - Non-parametric identification (Hu and Schennach, 2008).
  - Standard assumptions in the econometrics literature (Chen et al., 2011).
- Identification stems from an **exclusion restriction** that brings in the picture a particular type of instrument Z.
  - The variable Z can be arbitrarily correlated with errors.
- Identification rests upon **three key assumptions** (possibly conditional on household and area characteristics).
  - Additional mild regularity conditions are needed for identification.

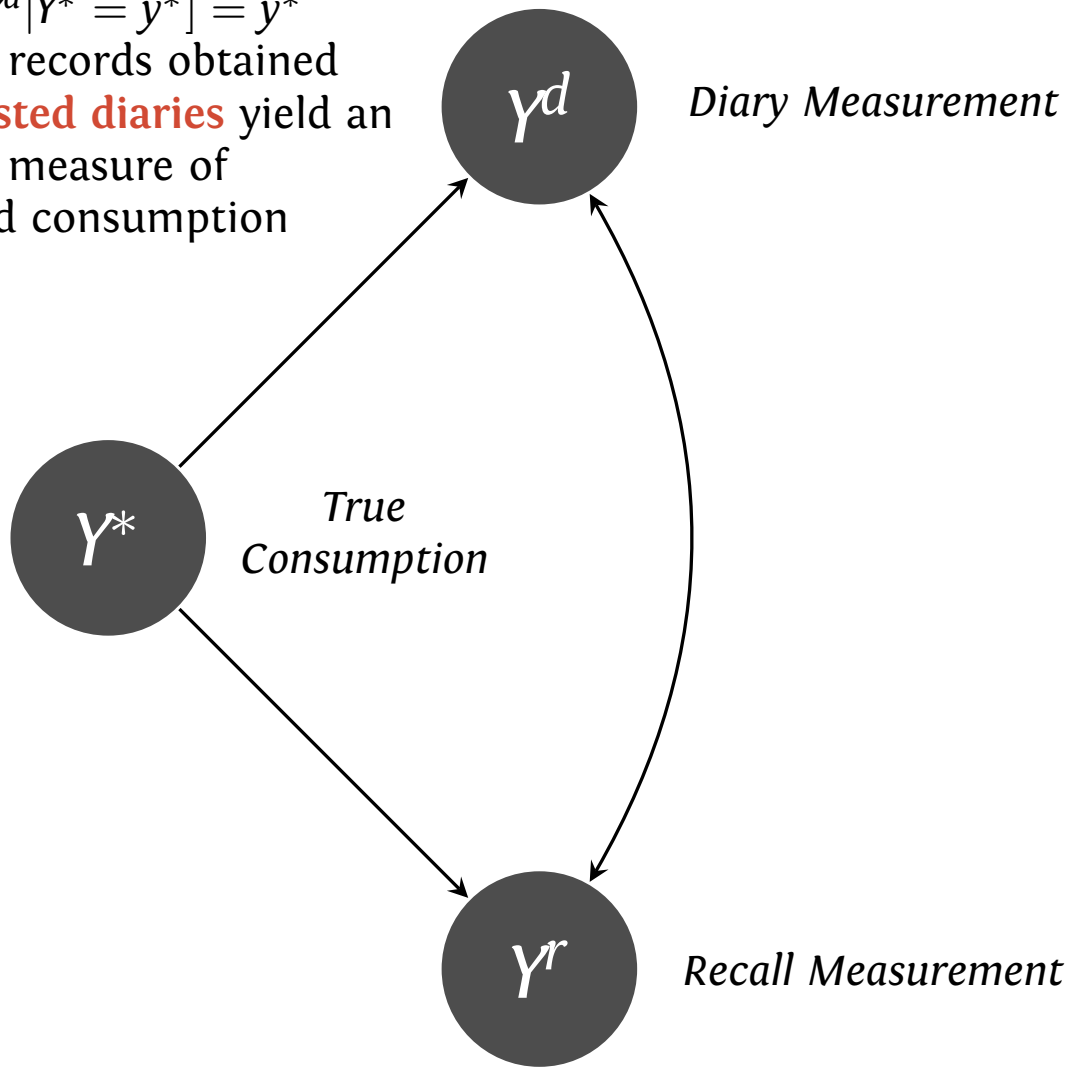
Assumptions in a Nutshell



Assumptions in a Nutshell

$$E[Y^d | Y^* = y^*] = y^*$$

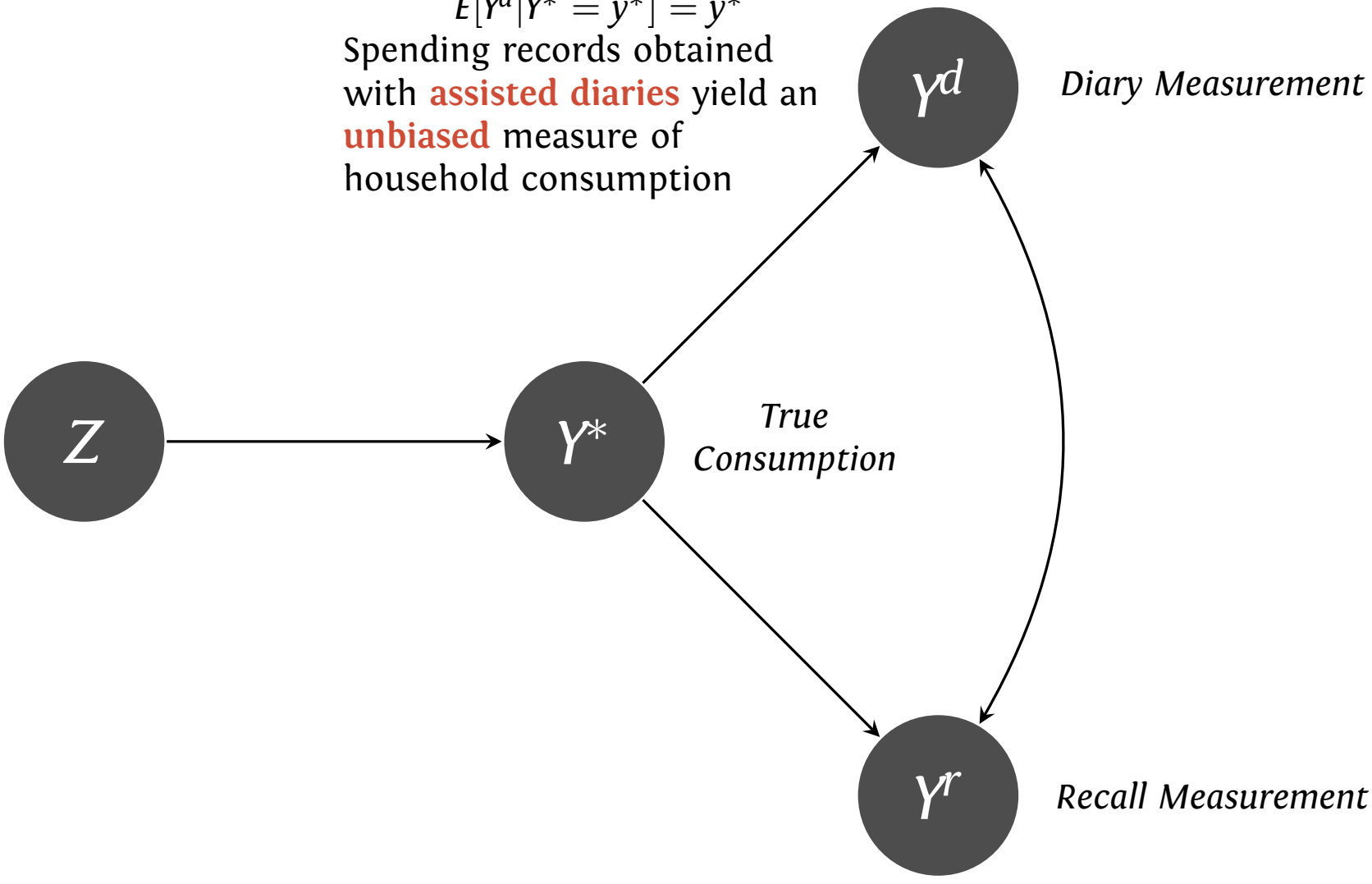
Spending records obtained with **assisted diaries** yield an **unbiased** measure of household consumption



Assumptions in a Nutshell

$$E[Y^d|Y^* = y^*] = y^*$$

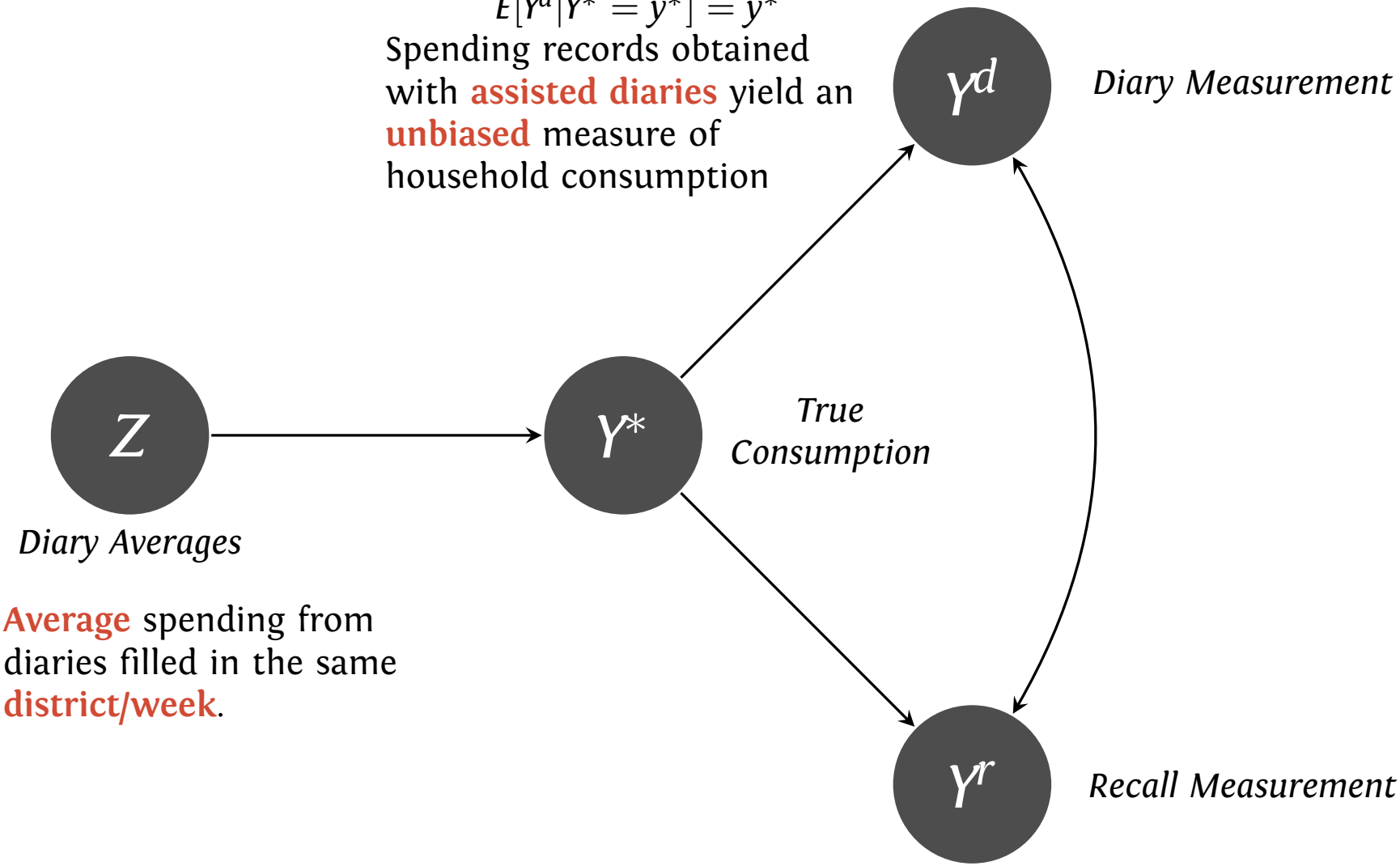
Spending records obtained with **assisted diaries** yield an **unbiased** measure of household consumption



Assumptions in a Nutshell

$E[Y^d|Y^* = y^*] = y^*$

Spending records obtained with **assisted diaries** yield an **unbiased** measure of household consumption



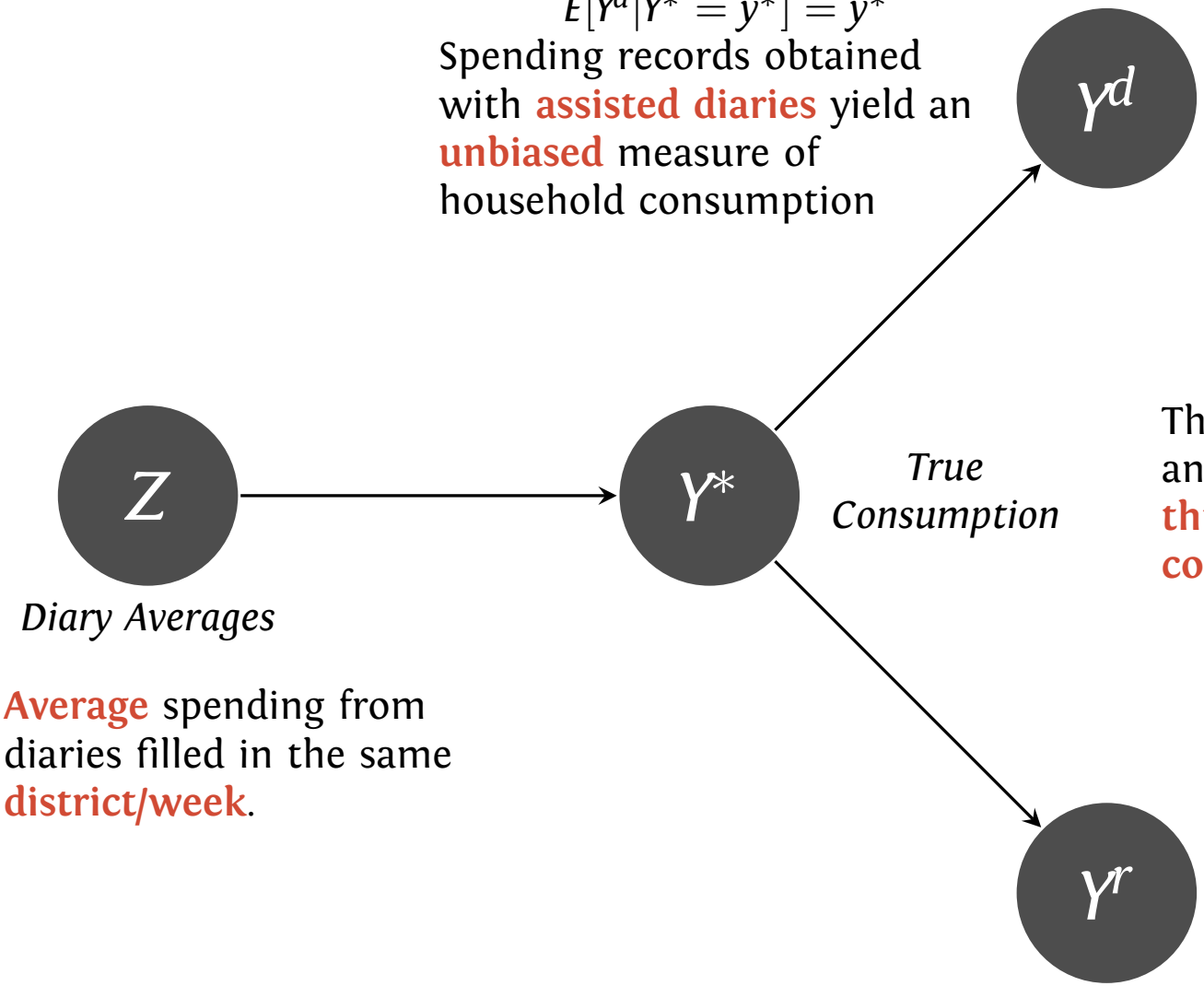
Diary Averages

Average spending from diaries filled in the same **district/week**.

Assumptions in a Nutshell

$$E[Y^d|Y^* = y^*] = y^*$$

Spending records obtained with **assisted diaries** yield an **unbiased** measure of household consumption



The correlation between diary and recall errors is **channeled through** the underlying **true consumption**

Average spending from diaries filled in the same **district/week**.

Estimation

- The three conditions above imply:

$$f_{Y^d Y^r|Z}[y^d, y^r|z] = \int \underbrace{f_{Y^d|Y^*}[y^d|y^*]}_{\text{diary errors}} \underbrace{f_{Y^r|Y^*}[y^r|y^*]}_{\text{recall errors}} \underbrace{f_{Y^*|Z}[y^*|z]}_{\text{true}} dy^*.$$

- **Non-parametric identification**: there exists a unique choice of distributions on the right-hand side that generates the observable distribution on the left-hand side (Hu and Schennach, 2008).

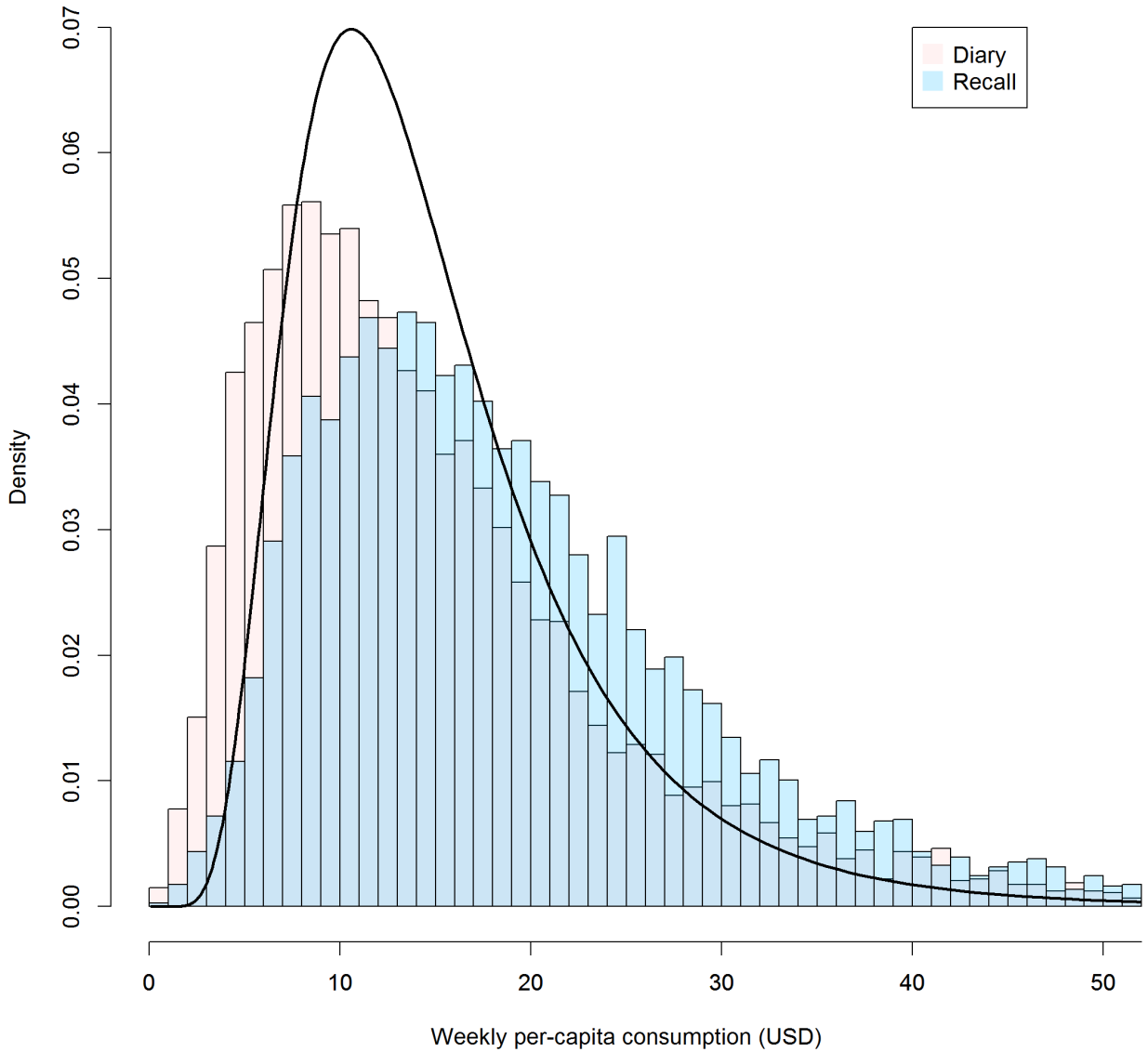
Estimation

- The three conditions above imply:

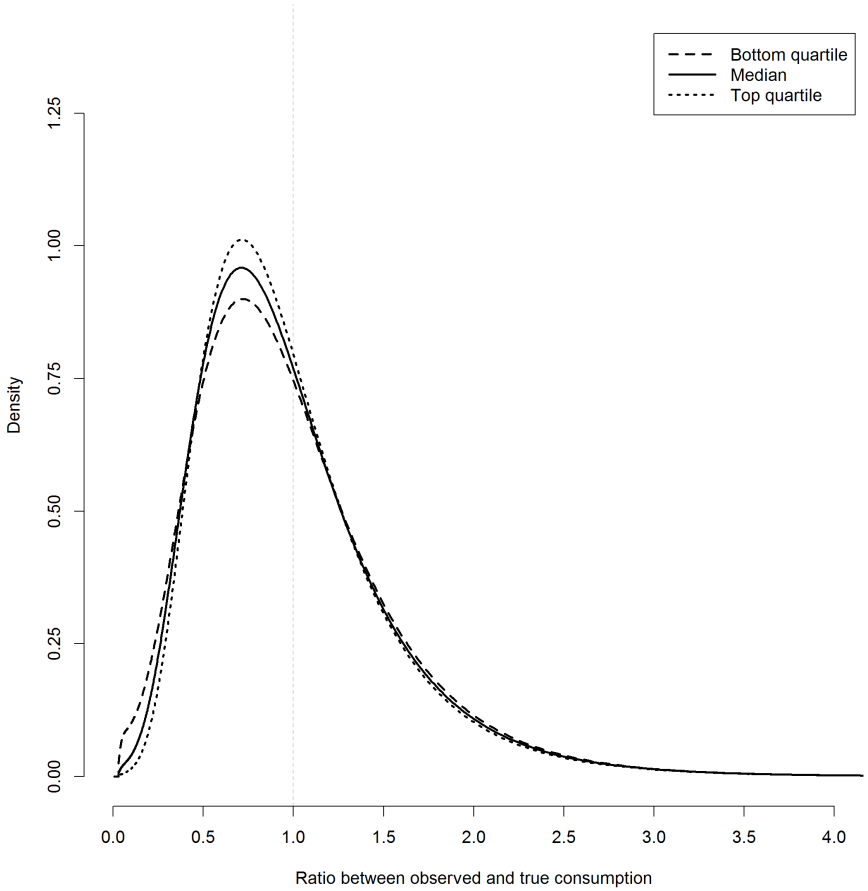
$$f_{Y^d Y^r|Z}[y^d, y^r|z] = \int \underbrace{f_{Y^d|Y^*}[y^d|y^*;\theta_d]}_{\text{diary errors}} \underbrace{f_{Y^r|Y^*}[y^r|y^*;\theta_r]}_{\text{recall errors}} \underbrace{f_{Y^*|Z}[y^*|z;\theta_y]}_{\text{true}} dy^*.$$

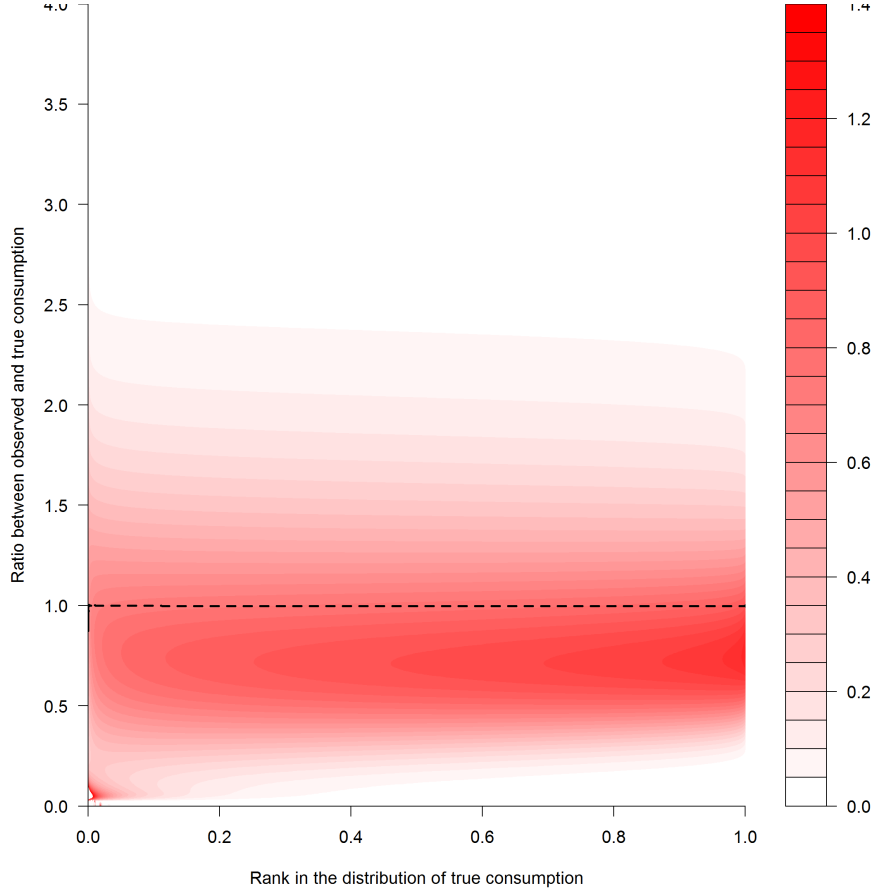
- **Non-parametric identification**: there exists a unique choice of distributions on the right-hand side that generates the observable distribution on the left-hand side (Hu and Schennach, 2008).
- Estimation: **sieve maximum-likelihood**; flexible specifications encompassing a rich family of distributions, yielding non-parametric estimates of the conditional densities on the right hand side.
- Use estimates of conditional densities to obtain **quantities of interest**, i.e. marginal distribution of true consumption ($f_{Y^*}[y^*]$).

Results: True Consumption



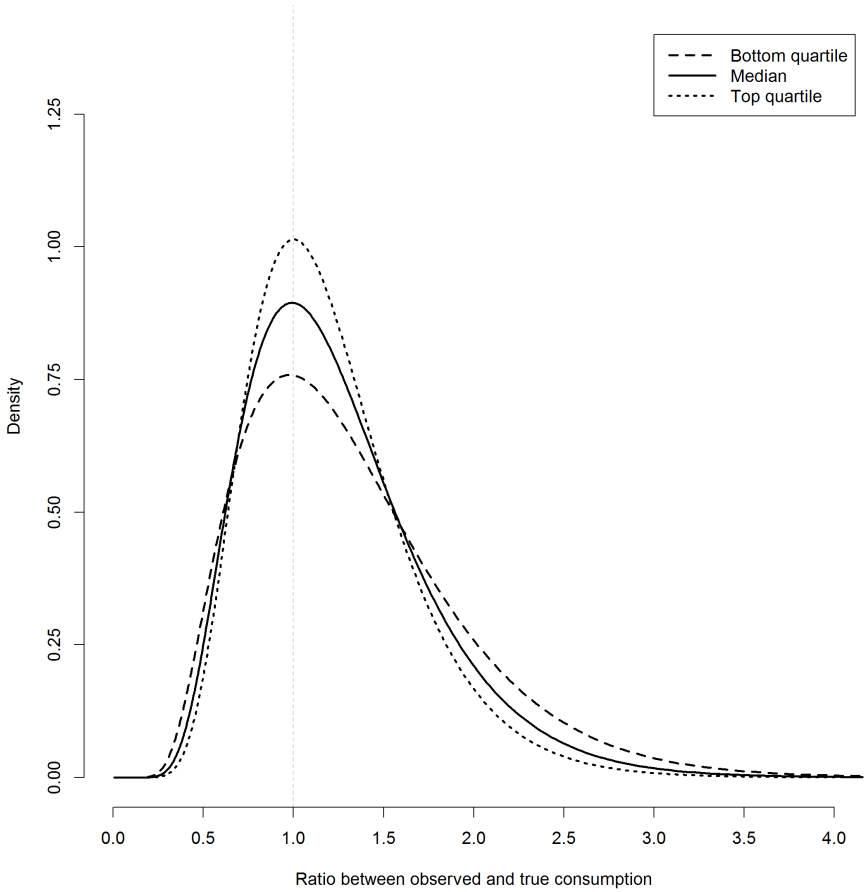
Diary Measurement Errors

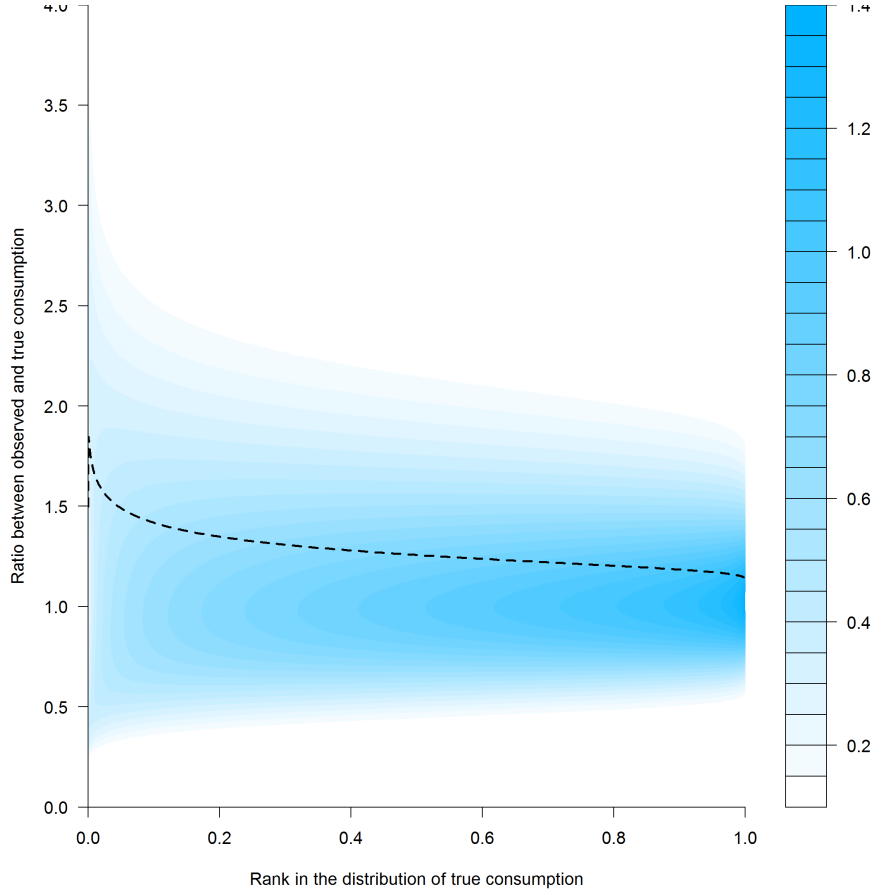




- Despite unbiasedness (i.e. $E[Y^d|Y^* = y^*] = y^*$) diaries yield **under-reported** consumption.
- Little differences across households with different levels of consumption.

Recall Measurement Errors

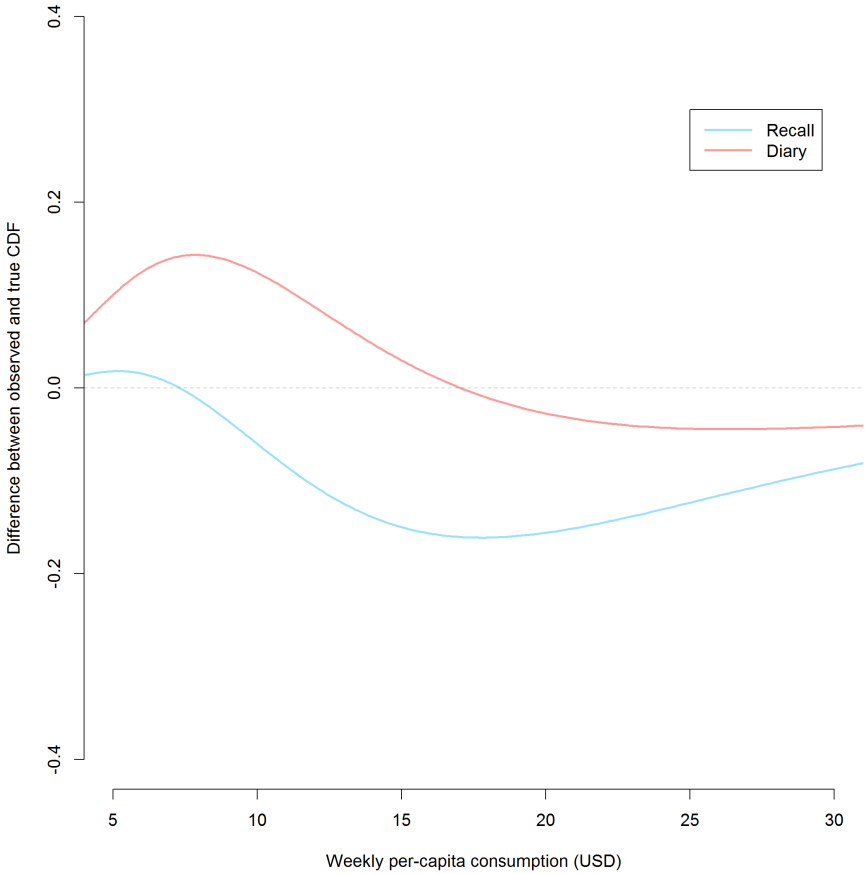




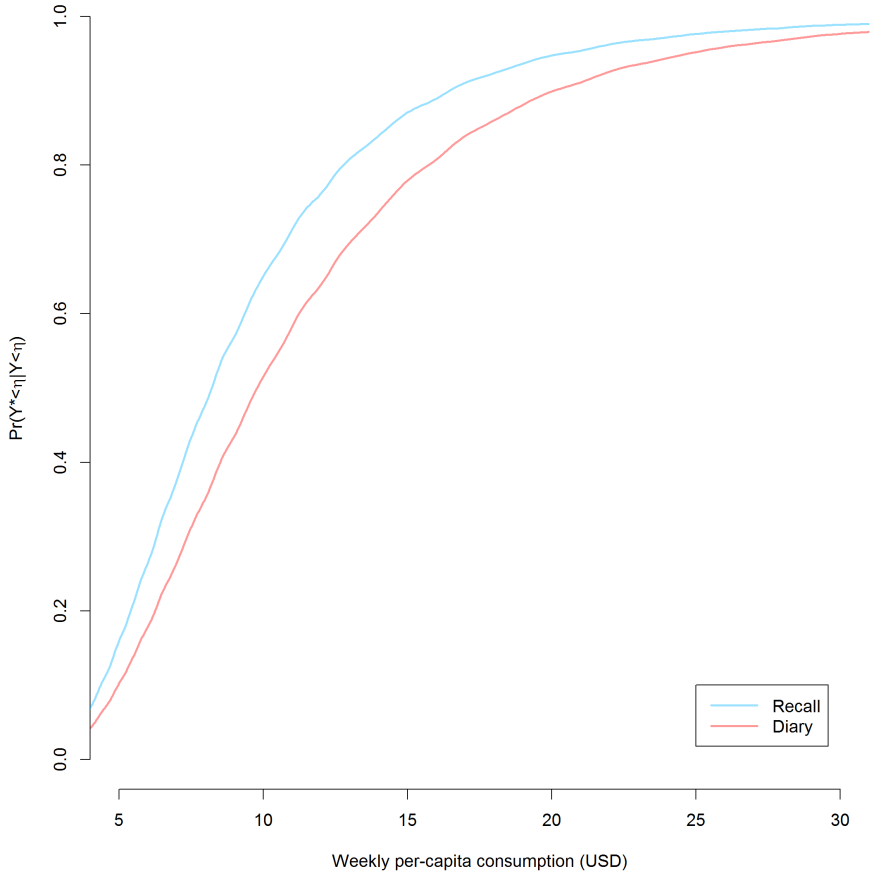
- Errors are **not classical**. Modal entry about right.
- **Smaller errors** for households with high levels of consumption.

Targeting

Poverty Measurement vs. True Poverty using different poverty lines



Share of poor household correctly classified



Optimal Assignment to Diary and Recall Interviews

- Consider a setting where households are **assigned a diary** with probability $p \in [0, 1]$.
- The **observed distribution** $F_Y(y; p)$ arising from this design is:

$$F_Y(y; p) = F_{Yd}(y)p + F_{Yr}(y)(1 - p).$$

- We are interested in the effects of this assignment on **functionals** of the distribution of observed consumption (e.g., the Gini coefficient): $\nu(F_Y(y; p))$.
- Knowledge of the true distribution of consumption $F_{Y^*}(y^*)$ allows to **compare**

$$\nu(F_Y(y; p)) \qquad \text{vs.} \qquad \nu(F_{Y^*}(y^*)),$$

at any given level of p.

Optimal Assignment to Diary and Recall Interviews

- Now consider a setting where, at each value y^* of Y^*, households are **assigned a diary** with probability $p(y^*) \in [0, 1]$.
- The share of survey participants filling out a diary is:

$$p \equiv \int p(y^*) dF_{Y^*}(y^*).$$

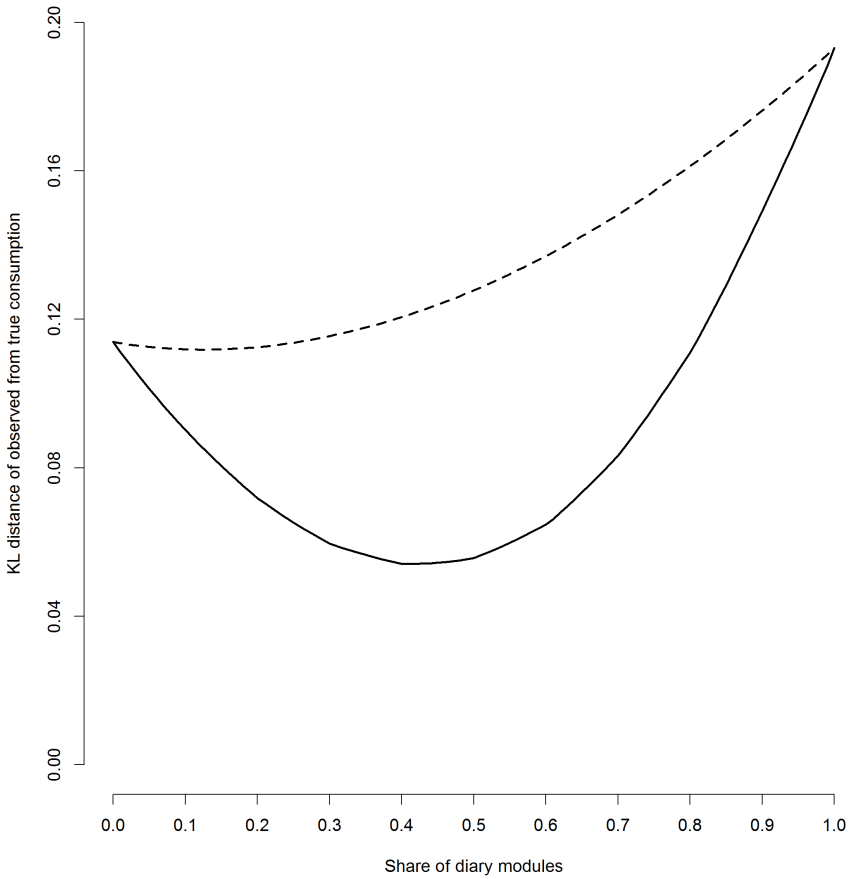
- The **observed distribution** $F_Y(y; p(y^*))$ arising from this design is:

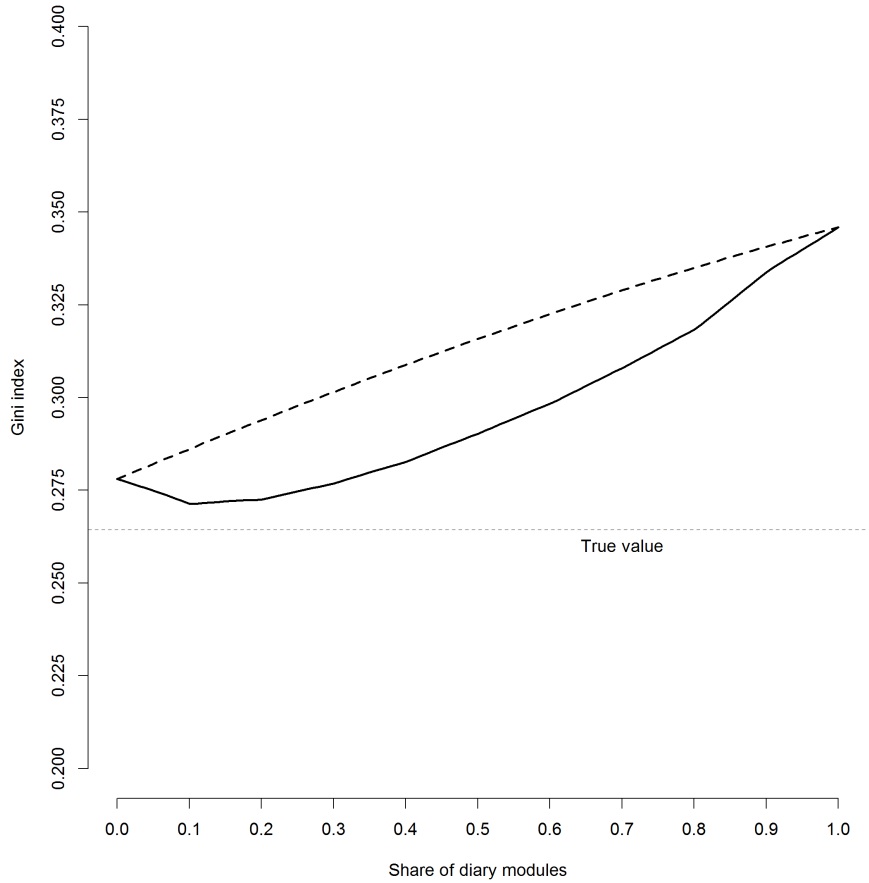
$$F_Y(y; p(y^*)) = \int \left[F_{Y^d|Y^*}(y|y^*) p(y^*) + F_{Y^r|Y^*}(y|y^*)(1 - p(y^*)) \right] dF_{Y^*}(y^*).$$

- For any given share of diaries p we can obtain the optimal assignment rule $p(y^*)$ which **minimizes** the difference

$$\nu \left(F_Y(y; p(y^*)) \right) - \nu \left(F_{Y^*}(y^*) \right).$$

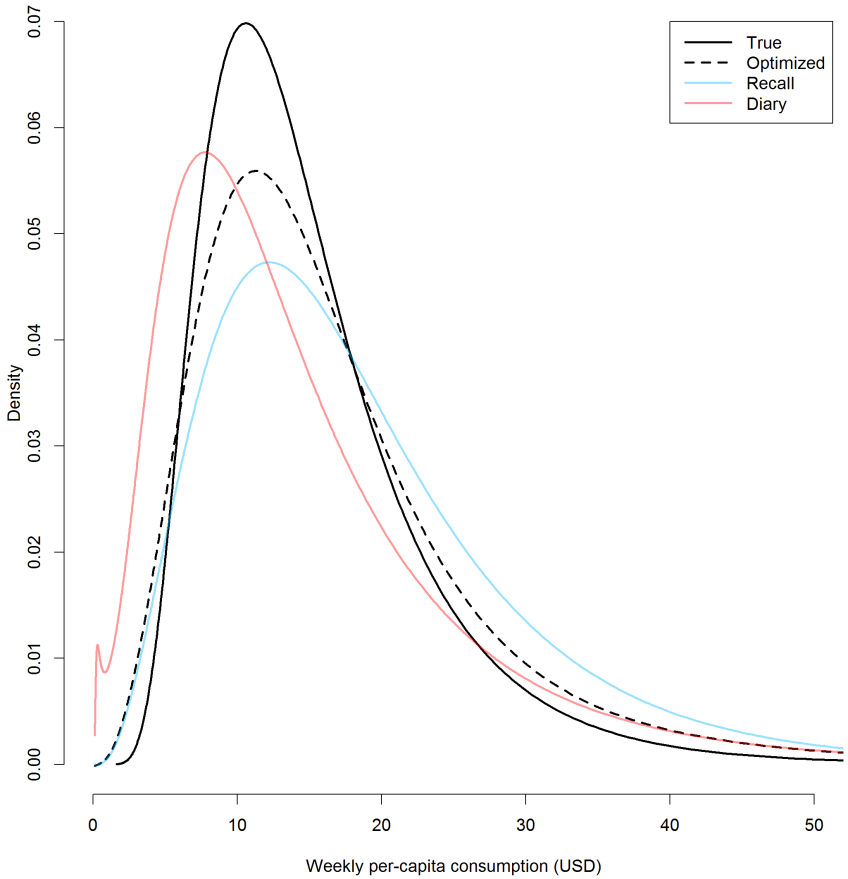
Optimal Mix Results

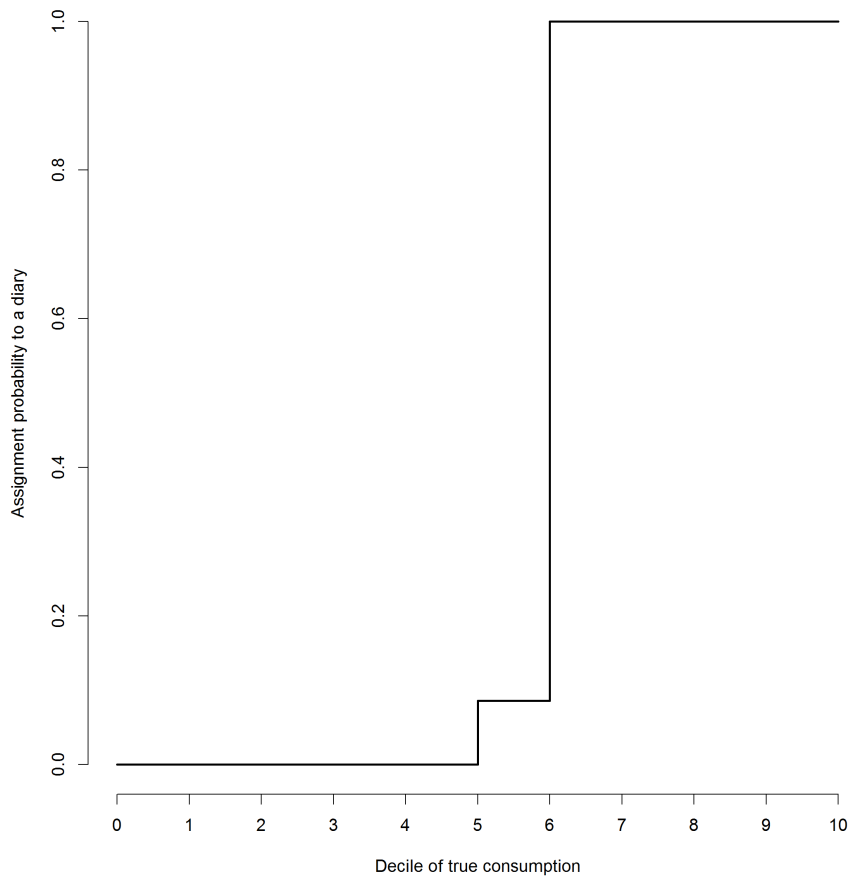




- Dashed line: **fixed** probability of being assigned to diary interview, i.e. $p(y^*) = p$.
- Solid line: probability of being assigned to diary interview **varying with Y^***.

Optimal Mix Results



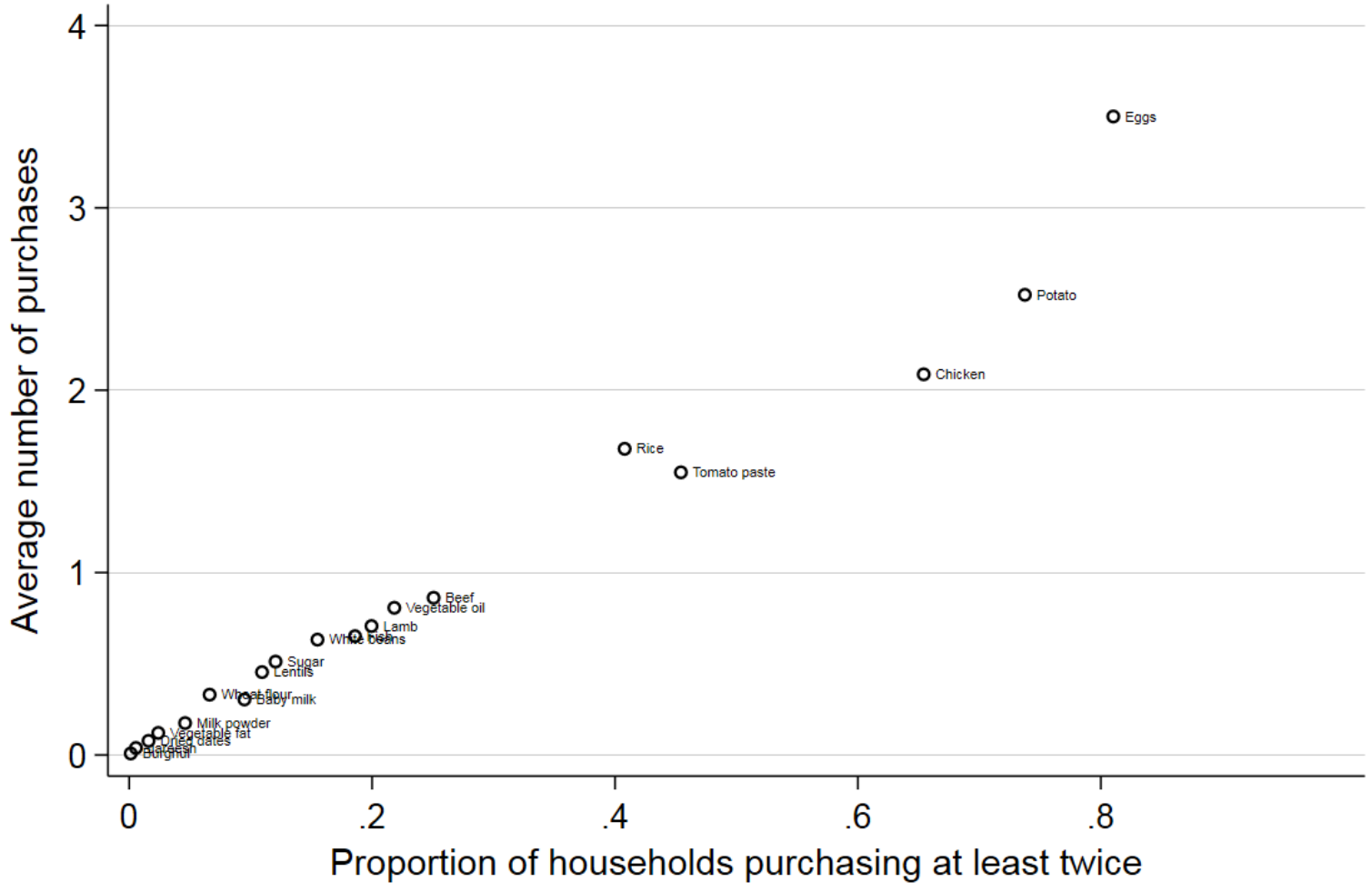


- Optimal mix allocates diaries to households on the **right tail** of the distribution of true consumption.
- In practice Y^* is unknown; a **feasible survey design** would allocate diaries based on a **proxy** for Y^*.

Conclusions

- **Little empirical support** for the idea that diaries yield data of **better quality** for measuring household welfare.
- Loss in accuracy in using recall questions to measure poverty is minimal compared to the **increasing costs** of using diaries.
- Even more so when **inequality** and **poverty measurement** is of interest.
- Diaries should collect information about **frequency** of consumption/purchase to correct for the potentially large errors.

Appendix - Items Frequency



Appendix - Balancing Tests

	Bimester					
	II	III	IV	V	VI	F-test
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Household characteristics						
Age	-0.040	-0.007	-0.033	-0.017	0.008	0.339
	(*0.026*)	(*0.025*)	(*0.025*)	(*0.026*)	(*0.025*)	
Education level	0.036	0.030	0.017	0.048	0.067**	0.294
	(*0.031*)	(*0.030*)	(*0.031*)	(*0.030*)	(*0.030*)	
Employed	-0.010	0.008	0.030	0.004	0.024	0.505
	(*0.023*)	(*0.023*)	(*0.023*)	(*0.023*)	(*0.023*)	
Panel B. Spending and prices						
Log expenditure	0.204***	0.098***	0.080***	0.109***	0.030	0.000
	(*0.022*)	(*0.021*)	(*0.021*)	(*0.022*)	(*0.022*)	
Log price index	0.026***	0.026***	-0.017***	0.008**	0.007**	0.000
	(*0.004*)	(*0.003*)	(*0.003*)	(*0.003*)	(*0.003*)	
Exp. share on rice	0.180***	0.202***	0.232***	0.025	0.027	0.000
	(*0.022*)	(*0.024*)	(*0.022*)	(*0.021*)	(*0.021*)	
Exp. share on potatoes	-0.146***	-0.162***	-0.245***	-0.135***	-0.059**	0.000
	(*0.027*)	(*0.027*)	(*0.027*)	(*0.029*)	(*0.029*)	
Exp. share on eggs	-0.137***	-0.078***	-0.184***	-0.184***	0.032	0.000
	(*0.025*)	(*0.026*)	(*0.025*)	(*0.026*)	(*0.027*)	
Exp. share on meat	0.012	0.108***	0.222***	0.361***	0.107***	0.000
	(*0.023*)	(*0.023*)	(*0.023*)	(*0.025*)	(*0.024*)	
Exp. share on fish	0.011	-0.047**	-0.088***	-0.088***	-0.039*	0.000
	(*0.022*)	(*0.021*)	(*0.022*)	(*0.021*)	(*0.022*)	

Appendix - Instrument Relevance

	Bimester					
	II	III	IV	V	VI	F-test
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Household characteristics						
Age	-0.040	-0.007	-0.033	-0.017	0.008	0.339
	(*0.026*)	(*0.025*)	(*0.025*)	(*0.026*)	(*0.025*)	
Education level	0.036	0.030	0.017	0.048	0.067**	0.294
	(*0.031*)	(*0.030*)	(*0.031*)	(*0.030*)	(*0.030*)	
Employed	-0.010	0.008	0.030	0.004	0.024	0.505
	(*0.023*)	(*0.023*)	(*0.023*)	(*0.023*)	(*0.023*)	
Panel B. Spending and prices						
Log expenditure	0.204***	0.098***	0.080***	0.109***	0.030	0.000
	(*0.022*)	(*0.021*)	(*0.021*)	(*0.022*)	(*0.022*)	
Log price index	0.026***	0.026***	-0.017***	0.008**	0.007**	0.000
	(*0.004*)	(*0.003*)	(*0.003*)	(*0.003*)	(*0.003*)	
Exp. share on rice	0.180***	0.202***	0.232***	0.025	0.027	0.000
	(*0.022*)	(*0.024*)	(*0.022*)	(*0.021*)	(*0.021*)	
Exp. share on potatoes	-0.146***	-0.162***	-0.245***	-0.135***	-0.059**	0.000
	(*0.027*)	(*0.027*)	(*0.027*)	(*0.029*)	(*0.029*)	
Exp. share on eggs	-0.137***	-0.078***	-0.184***	-0.184***	0.032	0.000
	(*0.025*)	(*0.026*)	(*0.025*)	(*0.026*)	(*0.027*)	
Exp. share on meat	0.012	0.108***	0.222***	0.361***	0.107***	0.000
	(*0.023*)	(*0.023*)	(*0.023*)	(*0.025*)	(*0.024*)	
Exp. share on fish	0.011	-0.047**	-0.088***	-0.088***	-0.039*	0.000
	(*0.022*)	(*0.021*)	(*0.022*)	(*0.021*)	(*0.022*)	