Consumption effects of subsidies in low-income households: Evidence from bivariate copula-based quantiles

Franziska Dorn* Simone Maxand**

* University of Duisburg-Essen
** European University Viadrina Frankfurt (Oder)

EAEPE Bilbao
September 2024

Project goals

- Bivariate energy-food poverty (richness and inequality) accounting for the interconnection of the consumption categories in all quantiles.

- Evaluating policies that target bidimensional poverty in energy and food consumption e.g., energy subsidies, food subsidies.
- or overconsumptions, e.g. or tax reforms (future).

- For this we need: best modeling techniques for bivariate quantiles
  - The whole distribution,
  - The drivers of the bivariate distribution,
  - Missing data at the upper tail (future).

Poverty Measurement

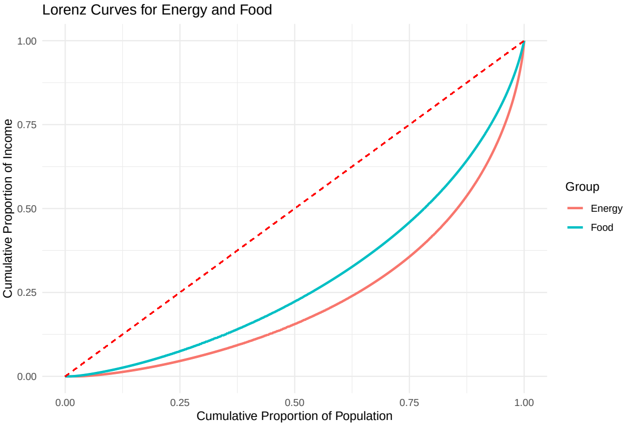
- Being vulnerable in multiple dimensions is more severe than being deficient in one (Pogge, 2002).
- Distributional aspects matter to understand intersectional aspects (Dorn et al., 2023).
- Multidimensional poverty measurement: Often indices or mean regression (Alkire and Foster, 2011; Alkire et al., 2015)
- Absolute vs. relative poverty in the US (Notten and Neubourg, 2007).
- Sustainability thresholds (Fanning et al., 2022).
- Affordability (Dogan et al., 2022)/ Investment constraints of low-income households for energy efficiency transformation (Cayla et al., 2011).

Data

2003-2019 US Panel Study of Income Dynamics (PSID)

- Couple and single adult households: 70118 observations
- Bivariate dependent vector variables:
  - Food expenditure: Share of household food expenditure in household income.
    single and couple households
  - Energy expenditure: Share of household expenditure on electricity and gas or other types of fuel in household income.
- The covariates (following, e.g., Dogan et al., 2022)
  - $race$: comprises seven categories
  - $gender$: is divided into a binary category
  - $heattype$: describes how the department is heated in 12 categories
  - $energsubd$: Dummy for government subsidy for energy expenditure
  - $foodsubd$: Dummy for government subsidy for food expenditure
  - edu_yr: measures the years of education
  - age: the years of life
  - $hhsize$: represents the number of household members
  - $hhtype$: the building type in seven categories

Energy and food inequality in the US



$\text{GINI}_{energy} = 0.535$ and $\text{GINI}_{food} = 0.415$ based on PSID data.

$\Rightarrow$ Measure inequality based on conditional top and bottom shares.

Energy subsidies in the US

Major energy subsidy program in/before 2019:
Low Income Home Energy Assistance Program (LIHEAP) and
Weatherization Assistance Program (WAP)

- Effect of LIHEAP on energy insecurity (Murray and Mills, 2014)
- LIHEAP as a response to energy poverty (Bednar and Reames, 2020)
- Eligibility to LIHEAP:
  - being at or below 150% of the federal poverty level (FPL) income guidelines; or
  - 60 percent of the state median income.

Bivariate Relative Poverty Line (BRPL)

From empirical distribution following Dorn et al. (2023) and Klein and Kneib (2020):

- Joint cumulative distribution function of income (q_1) and leisure (q_2):

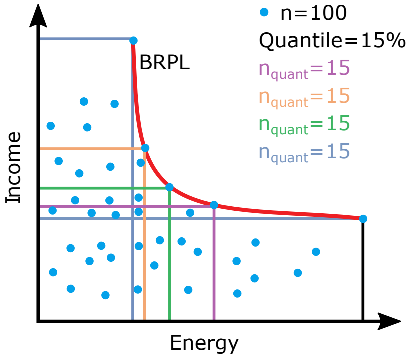
$$F_{1,2}(q_1, q_2)$$

- Bivariate poverty line is defined by fixing a quantile level:

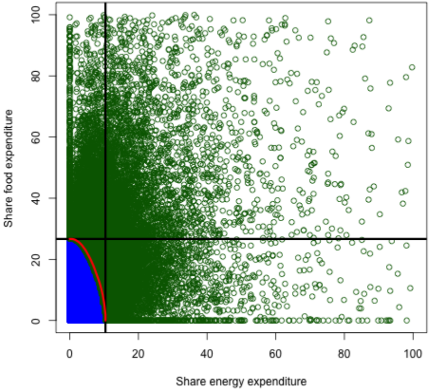
$$\tau \in [0, 1]$$

- The contour line (red) is determined by:

$$F_{1,2}(q_1, q_2) = \tau = 15\%$$



BRPL Empirics

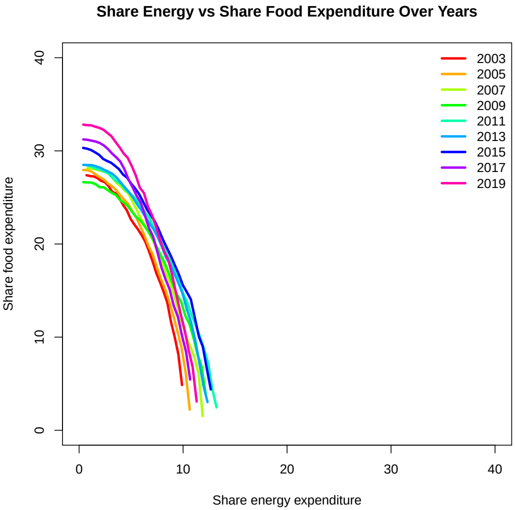


BRPL (red line) set to the 88% quantile food and energy share of income

Food: 26.7% and Energy: 10.4% of income

Poverty	Energy	Food	Absolute	Relative
Total	7.54	7.61	5.25	6.51

BRPL over time



Distributional copula regression

as in Dorn et al. (2024)

Systematically choose between alternative copulas based on AIC, QQ-plots or predictive risk with boosting as in Hans et al. (2023). Here: DAGUM distribution and normal copula.

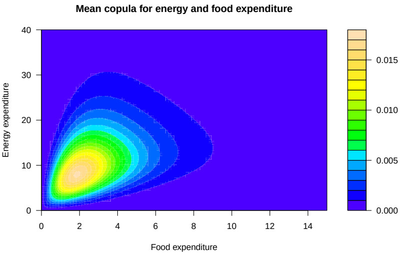
Response variables $Y_1 = energy$ and $Y_2 = food$ and bivariate distribution $D = F_{1,2}$ is defined by

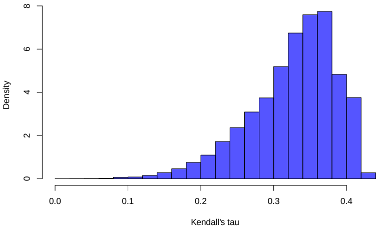
$$\begin{pmatrix} energy \\ food \end{pmatrix} \sim D(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7)$$

with parameter θ_j defined by a predictor $\theta_j(\mathbf{z}) = h_j(\eta^{\theta_j})$ with

$$\begin{aligned} \eta^{\theta_j} = \beta_0^{\theta_j} + \beta_2^{\theta_j} race + \beta_3^{\theta_j} gender + \beta_4^{\theta_j} heat\ type + \beta_5^{\theta_j} govsubd \\ + \beta_6^{\theta_j} foodsubd + \beta_7^{\theta_j} educ\_yr + \beta_8^{\theta_j} hhsize + \beta_9^{\theta_j} hhtype + \beta_{10}^{\theta_j} age \end{aligned}$$

Copula regression results





Bivariate conditional poverty risks:

	Kendall's tau	ThEnergy	ThFood	ThEnergyFood
Whole sample	0.33	0.11	0.14	0.20
White	0.33	0.10	0.13	0.19
Black	0.32	0.13	0.16	0.23
Energy subsidies	0.33	0.11	0.14	0.20
No energy subsidies	0.38	0.05	0.12	0.15

Conclusions and outlook

- 5.25% of US households are absolute and 6.51% of US households are relatively energy and income poor.
- Black people are more likely living in households below the joint threshold.
- Lower relation between food and energy expenditure for households with energy subsidies.

- Future work
  - Bivariate copula-based conditional quantiles
    - Identifying quantile-specific (causal) effects
    - Work in instrumental variable estimation and figure out good identification strategy.
  - Work out projections for the implementation of the Inflation Reduction Act.

References

Abdous, B. and R. Theodorescu (1992). Note on the spatial quantile of a random vector. *Statistics & Probability Letters 13*(4), 333–336.

Alkire, S. and J. Foster (2011). Understandings and misunderstandings of multidimensional poverty measurement. *The Journal of Economic Inequality 9*(2), 289–314.

Alkire, S., J. E. Foster, S. Seth, M. E. Santo, and J. M. Roche (2015). Multidimensional poverty measurement and analysis. *Oxford Poverty & Human Development Initiative OPHI Working Paper 84*.

Bednar, D. and T. Reames (2020). Recognition of and response to energy poverty in the united states. *Nat Energy 5*, 432–439.

Bricker, J., A. Henriques, J. Krimmel, and J. Sabelhaus (2016). Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity*, 261–312.

Brunori, P., P. Salas-Rojo, and P. Verme (2022). Estimating Inequality with Missing Incomes. GLO Discussion Paper Series 1138, Global Labor Organization (GLO).

Camehl, A., D. Fok, and K. Gruber (2022). Multivariate quantile regression using superlevel sets of conditional densities. Tinbergen Institute Discussion Papers 22-094/III, Tinbergen Institute.

Carlier, G., V. Chernozhukov, and A. Galichon (2017). Vector quantile regression beyond the specified case. *Journal of Multivariate Analysis 161*, 96–102.

Cayla, J.-M., N. Maizi, and C. Marchand (2011). The role of income in energy consumption behaviour: Evidence from french households data. *Energy Policy 39*(12), 7874–7883. Clean Cooking Fuels and Technologies in Developing Economies.

Dogan, E., M. Madaleno, R. Inglesi-Lotz, and D. Taskin (2022). Race and energy poverty: Evidence from african-american households. *Energy Economics 108*, 105908.

Dorn, F., S. Maxand, and T. Kneib (2024). The nonlinear dependence of income inequality and carbon emissions: Potentials for a sustainable future. *Ecological Economics*.

Dorn, F., R. Radice, G. Marra, and T. Kneib (2023). A bivariate relative poverty line for leisure time and income poverty: Detecting intersectional differences using distributional copulas. *Review of Income and Wealth*.

Fanning, A., D. Oâ€™Neill, and J. Hickel (2022). The social shortfall and ecological overshoot of nations. *Nature Sustainability 5*, 26â€“36.

Goldstein, B., D. Gounaridis, and J. P. Newell (2020). The carbon footprint of household energy use in the united states. *Proceedings of the National Academy of Sciences 117*(32), 19122–19130.

Hallin, M., D. Paindaveine, and M. Å iman (2010). Multivariate quantiles and multiple-output regression quantiles: From L1 optimization to halfspace depth. *The Annals of Statistics 38*(2), 635 – 669.

Hans, N., N. Klein, F. Faschingbauer, M. Schneider, and A. Mayr (2023). Boosting distributional copula regression. *Biometrics 79*(3), 2298–2310.

Harold, J., J. Cullinan, and S. Lyons (2017). The income elasticity of household energy demand: a quantile regression analysis. *Applied Economics 49*(54), 5570–5578.

Heathcote, J., F. Perri, and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the united

Definitions of Poverty, Richness and Inequality

	Energy expenditures	Income
Poverty	Affordability (Dogan et al., 2022) Investment constraints of low-income households for energy efficiency transformation (Cayla et al., 2011)	Absolute vs relative poverty in the US (Notten and Neubourg, 2007)
Richness	Energy footprint and over-consumption (Goldstein et al., 2020)	Rich vs wealthy (Bricker et al., 2016), Measurement issues with survey data (Piketty et al., 2022; Pfeffer et al., 2016).
Inequality	Lorenz curve (Oswald et al., 2020), GINI index (Jacobson et al., 2005)	In the US (Heathcote et al., 2010), sustainability thresholds (Fanning et al., 2022).

The Energy-Income Nexus

- Income elasticity of energy use in quantiles (Kaza, 2010; Harold et al., 2017).
- Linn et al. (2022) negative relation between income inequality and energy use in the US.

Bivariate copula-based conditional quantiles

- Bivariate quantiles are not uniquely defined, e.g., via statistical depths, calculating vector-based or spatial quantiles (see, e.g., Hallin et al., 2010; Carlier et al., 2017; Abdous and Theodorescu, 1992).

- We are primarily interested in the upper-right or the lower-left corner. $\rightarrow$ See Klein and Kneib (2020) and Tepegjozova and Czado (2022).

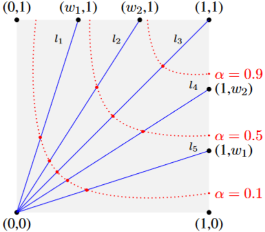
- Our approach is similar to Tepegjozova and Czado (2022):
  - Relate bivariate distribution to covariates via distributional copula regression.
  - Define quantiles for conditional copula $C_{V_1,V_2|\boldsymbol{X}}(v_1, v_2|\boldsymbol{x})$ by

$$Q_{\alpha}^{Y|X}(x) = \{(F_{Y_1}(y_1), F_{Y_2}(y_2)) \in [0,1]^2; C_{V_1,V_2|\boldsymbol{X}}(v_1, v_2|\boldsymbol{x}) = \alpha\},$$

  - Determine $Q_{\alpha}^{Y|X}(x)$ by line search algorithm.

Derivation of the conditional quantiles $Q_{\alpha}^{Y|X}(x)$

1. Choose a combination of covariates $\boldsymbol{x}_0 \in \mathbb{R}^K$ and store conditional distribution parameters $\hat{\theta}_1(\boldsymbol{x}_0), \ldots, \hat{\theta}_7(\boldsymbol{x}_0)$.
2. We determine values $(v_1^*, v_2^*) \in [0,1]^2$ for which the condition $C_{V_1,V_2|X_0}(v_1, v_2|\boldsymbol{x}_0) = \alpha$ holds with $C_{V_1,V_2|X_0}$ defined by $\hat{\theta}_7(\boldsymbol{x}_0)$. Follow a line search algorithm to determine $(v_1^*, v_2^*) \in [0,1]^2$.



3. Values (v_1^*, v_2^*) are transformed to the original data space by the inverse conditional distributions $F^{-1}_{Y_1|\mathbf{x}_0}(v_1^*)$ and $F^{-1}_{Y_2|\mathbf{x}_0}(v_2^*)$. We obtain $Y^* = (y_1^*, y_2^*)$ which are on the quantile line $Q_{\alpha}^{Y|X}(\boldsymbol{x}_0)$.

Quantiles from the lower left for simulated data

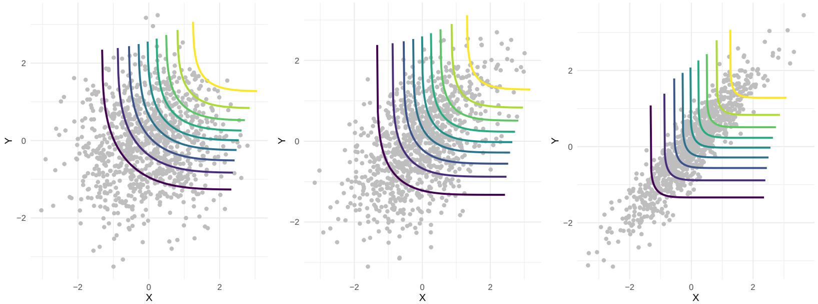


Figure 1: Simulated standard normally distributed data with $n = 1000$ and dependence parameters $\rho = 0.1, 0.5, 0.9$ (from left to right) and the calculated 10% to 90% lower-left quantiles when regressing on a constant.

Quantiles from the upper right for simulated data

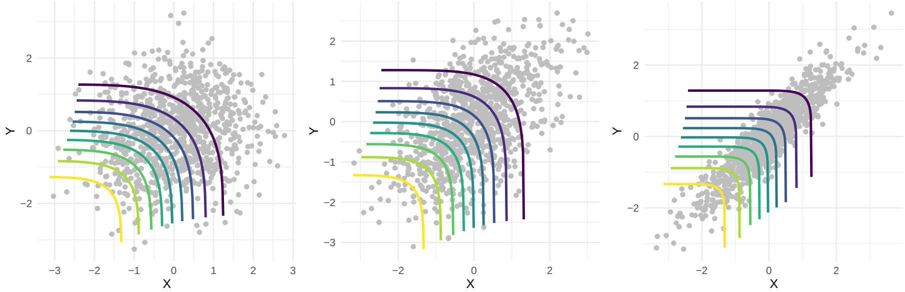


Figure 2: Simulated standard normally distributed data with $n = 1000$ and dependence parameters $\rho = 0.1, 0.5, 0.9$ (from left to right) and the calculated 10% to 90% upper-right quantiles obtained by the described method when regressing on a constant.

Identifying quantile-specific (causal) effects

Similar to Camehl et al. (2022); Sanchez et al. (2020), quantile-specific marginal effects of covariates:

$$\beta^j(\alpha|\mathbf{x}) = dist(Q_\alpha^{Y|X}(x + \Delta_k), Q_\alpha^{Y|X}(x))$$

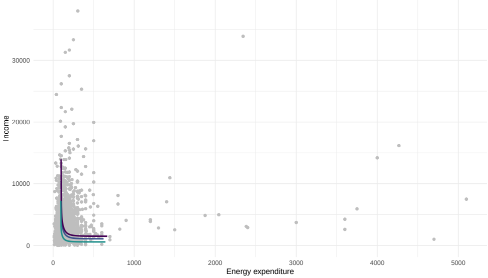
Marginal effect on one of the response variables by, e.g.,

$$\beta_1^j(\alpha|y_2, \mathbf{x}) = dist(Q_\alpha^{Y_1|Y_2}(x + \Delta_k), Q_\alpha^{Y_1|Y_2}(x)).$$

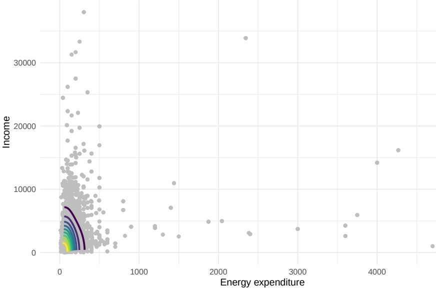
Causal effects by instrumental variable estimation: see two-step GAMLSS as in (Sanchez et al., 2020).

First results

The 10% quantile conditional on the mean population (in black), on the persons eligible for energy subsidy (in blue) and those who receive energy subsidies (in green):



The mean conditional bivariate quantiles:



Future work on this project

- Modelling unreported observations in high-income groups:
  - Top income biases on the measurement of inequality in the US (Hlasny and Verme, 2022; Bricker et al., 2016).
  - Brunori et al. (2022): the unreported income rich are best modeled by a Pareto distribution.
  - Quantile regression can be combined with a Pareto distribution (this is done via gradient boosting, for instance, in Velthoen et al., 2023)

- Work in instrumental variable estimation and figure out good identification strategy.
  - Policies in the US on food and energy subsidies?

- Extend data set to panel data in order to identify effects of tax reforms, e.g., by using event study analysis.

- Work out projections for the implementation of the Inflation Reduction Act.