The Bright Side of Heteroskedasticity: Measuring Quarterly Economic Growth from Outer Space

Robert C. M. Beyer   Yingyao Hu   Jiaxiong Yao

IMF JHU IMF

ES session in honor of Tom Wansbeek
IPDC 2026

Outline

Motivation

- GDP data suffer from significant gaps:
  - $\sim$1/3 of countries lack quarterly GDP figures
  - Substantial revisions, reporting lags, no subnational detail
- Satellite nighttime lights (NTL) are a compelling proxy for economic activity
  - Since Chen and Nordhaus (2011) and Henderson, Storeygard, and Weil (2012)
  - Used in: growth, poverty, trade, regional development, political economy, etc.
- Key component: the **elasticity** between NTL and economic activity
  - Needed to convert NTL changes into GDP changes
  - Both NTL and GDP contain measurement errors $\Rightarrow$ under-identification

Contributions

1. **New identification strategy** through heteroskedasticity
   - No assumptions on GDP data quality
   - Applicable to other satellite data
2. **Latest NTL data** (VIIRS) at annual and quarterly frequencies
   - Annual: 181 countries, 1993–2023 (simulated VIIRS)
   - Quarterly: 107 countries, 2013Q2–2025Q2
3. **Key findings:**
   - Elasticity ≈ 1.52 (annual), ≈ 2.25 (quarterly)
   - Stable across country groups and sectoral composition
   - GDP signal ratios: 96% (AEs) to 56% (LIDCs) annually

The Problem: Errors on Both Sides

NTL growth and GDP growth are both measured with error:

$$z_{i,t} = z^*_{i,t} + \varepsilon^z_{i,t} \qquad \text{(NTL growth)}$$
$$y_{i,t} = y^*_{i,t} + \varepsilon^y_{i,t} \qquad \text{(GDP growth)}$$
$$z^*_{i,t} = \beta\, y^*_{i,t} + \eta_{i,t} \qquad \text{(Structural relationship)}$$

- β = elasticity of NTL with respect to GDP
- All errors are classical (zero mean, finite variance)
- Growth rates = first difference of log levels

Consolidating: $z_{i,t} = \beta\, y^*_{i,t} + \underbrace{(\varepsilon^z_{i,t} + \eta_{i,t})}_{\mu^z_{i,t}}$

Under-Identification from Second Moments

Taking second moments:

$$\mathrm{var}(y) = \mathrm{var}(y^*) + \mathrm{var}(\varepsilon^y) \tag{S1}$$

$$\mathrm{var}(z) = \beta^2 \,\mathrm{var}(y^*) + \mathrm{var}(\mu^z) \tag{S2}$$

$$\mathrm{cov}(y, z) = \beta \,\mathrm{var}(y^*) \tag{S3}$$

3 equations, 4 unknowns: β, $\mathrm{var}(y^*)$, $\mathrm{var}(\varepsilon^y)$, $\mathrm{var}(\mu^z)$

$\Rightarrow$ System is **not identified** without additional information

Previous Approaches to Identification

1. **Assume signal-to-noise ratio** (Henderson, Storeygard, and Weil, 2012)
   - Assume $\text{Sig}_y = 90\%$ for good-data countries; derive the rest
   - Requires an external assumption about GDP data quality
2. **Impose structure on measurement errors** (Hu and Yao, 2022)
   - Link NTL errors to location; GDP errors to statistical capacity
   - Requires complex and flexible error specifications
3. **Add an additional GDP signal** (Civelli, Gaduh, and Yousuf, 2024)
   - Use a third data source to provide extra moment conditions
   - Requires availability of such a signal

This Paper

Identification through **heteroskedasticity** — grounded in the satellite data-generating process, no GDP assumptions needed

Key Insight: NTL Errors are Heteroskedastic

NTL $=$ sum of pixel lights averaged over space and time

Each pixel has idiosyncratic measurement errors. By the law of large numbers:

1. **Larger countries**: more pixels $\Rightarrow$ errors average out spatially
2. **More observations**: more daily obs $\Rightarrow$ errors average out temporally
3. **Brighter regions**: higher signal $\Rightarrow$ higher signal-to-noise ratio

Result (see Appendix for derivation)

$$\text{var}(\varepsilon^z_{i,t}) \approx \frac{2\sigma_z^2}{N_i \cdot N_i^{obs} \cdot (\mathrm{E}z_i)^2}$$

where σ_z^2 $=$ pixel-level error variance, $\mathrm{E}z_i$ $=$ average NTL per pixel

Derivation Sketch

Step 1: Temporal averaging at pixel level

$\varepsilon_{p,t} = \frac{1}{N_i^{obs}} \sum_{d=1}^{N_i^{obs}} \varepsilon_{p,d} \Rightarrow \mathrm{var}(\varepsilon_{p,t}) = \sigma_z^2 / N_i^{obs}$

Step 2: Spatial aggregation to country level

$\mathrm{var}(\varepsilon_{i,t}) = N_i \cdot \sigma_z^2 / N_i^{obs}$

Step 3: Growth rate (log difference) via Taylor expansion

$z_{i,t} \approx z_{i,t}^* + \frac{\varepsilon_{i,t}}{Z_{i,t}^*} - \frac{\varepsilon_{i,t-1}}{Z_{i,t-1}^*}$

Since $Z_{i,t}^* = \bar{z}_i \cdot N_i \approx \mathrm{E}z_i \cdot N_i$:

$$\mathrm{var}(\varepsilon_{i,t}^z) \approx \frac{2\,\mathrm{var}(\varepsilon_{i,t})}{(\mathrm{E}z_i N_i)^2} = \frac{2\sigma_z^2}{N_i \cdot N_i^{obs} \cdot (\mathrm{E}z_i)^2}$$

Identification through Heteroskedasticity

Replace pooled var(z) with **country-specific** variance:

$$\text{var}_i(z) = \beta^2 \, \text{var}(y^*) + \frac{2\sigma_z^2}{N_i \cdot N_i^{obs} \cdot (\text{E}z_i)^2} + \text{var}(\eta^z)$$

System: equations (S1), (S3), and the above for each country i

5 unknowns: β, var(y^*), var(ε^y), σ_z^2, var(η^z)

One equation per country $\Rightarrow$ **identified with** ≥ 3 **countries**

Comparison with Rigobon (2003)

Rigobon: all structural shocks are uncorrelated and heteroskedastic across two regimes. Here: the structural shocks are correlated and only NTL measurement errors are heteroskedastic.

Dealing with Quarterly Data

Challenge: Strong seasonality in NTL

- Number of observations varies across quarters (late sunsets in summer)
- Light-generating activities display seasonal patterns

Solution: Use **year-on-year** growth rates and quarter-specific moments:

$$\mathrm{var}_{i,q}(z) = \beta^2 \,\mathrm{var}(y^*) + \frac{2\sigma_z^2}{N_i \cdot N_{i,q}^{obs} \cdot (\mathrm{E}z_{i,q})^2} + \mathrm{var}(\eta^z)$$

where $q = 1, 2, 3, 4$

This addresses both the varying observation frequency and seasonal activity patterns.

Why is the Elasticity Important?

Application 1: Policy Analysis (Diff-in-Diff)

$$z_{i,t} = \psi_0 + \psi_1 \text{Treat}_i + \psi_2 \text{Post}_t + \psi_3 (\text{Treat}_i \times \text{Post}_t) + \epsilon_{i,t}$$

- ψ_3 captures impact on NTL growth
- Impact on GDP growth $= \hat{\psi}_3 / \hat{\beta}$ (requires structural elasticity)

Application 2: Predicting GDP

$$y_{i,t} = \gamma\, z_{i,t} + e_{i,t}$$

- plim $\gamma = \frac{1}{\beta}\left(\frac{\beta^2 \,\text{var}(y^*)}{\beta^2 \,\text{var}(y^*) + \text{var}(\mu^z)}\right)$
- Predictive coefficient $=$ inverse elasticity $\times$ signal ratio of NTL

Light-Adjusted GDP Measure

Combine official GDP and NTL-predicted GDP:

$$\bar{y} = (1 - \lambda^*)\, y + \lambda^*\, \tilde{y}, \quad \tilde{y} = \gamma\, z$$

Optimal weight minimizes expected MSE of true GDP prediction:

$$\lambda^* = \frac{\mathrm{var}(\varepsilon^y)}{\mathrm{var}(\varepsilon^y) + (\gamma\beta - 1)^2\, \mathrm{var}(y^*) + \gamma^2\, \mathrm{var}(\mu^z)}$$

- Larger GDP measurement error $\Rightarrow$ higher weight on NTL
- All components identified by our framework
- Optimal λ^*: 0.38 (annual), 0.63 (quarterly)

Data Sources

Nighttime Lights

- DMSP-OLS: 1992–2013 (1st gen)
- VIIRS: 2012–present (2nd gen)
  - Higher resolution
  - Improved low-light imaging
  - Monthly frequency
- Simulated VIIRS (SVNL): 1992–2023
  - Chen, Wang, Zhang, Shen, and Chen (2024)
  - Deep learning transformation

Economic Data

- Annual GDP: IMF WEO
- Quarterly GDP: Haver Analytics
- Country groups: IMF classification
  - AEs, EMs, LIDCs
- GDP per capita: World Bank WDI
- Sectoral shares: World Bank WDI

Summary Statistics: SVNL Annual Dataset

	Mean	p1	p10	p25	p50	p75	p90	p99
AEs								
NTL growth	3.2	-67.4	-18.2	-6.6	1.3	11.1	24.8	92.4
GDP growth	2.5	-8.9	-1.1	1.1	2.6	4.2	6.3	10.8
avg. # of obs per year	87.2	47.1	54.0	62.3	73.4	109.3	143.6	178.4
area (thousand sq. km)	1184.7	0.07	3.3	46.3	98.3	458.8	790.9	16864.8
avg. pixel NTL (radiance)	2.3	0.01	0.09	0.3	0.7	1.9	4.6	32.5
EMs								
NTL growth	6.4	-61.1	-13.7	-3.6	5.0	14.6	28.9	84.1
GDP growth	3.3	-15.3	-2.0	1.2	3.6	5.7	8.1	18.6
avg. # of obs per year	109.2	30.9	55.8	77.7	102.2	139.4	171.8	193.9
area (thousand sq. km)	936.8	0.04	0.4	10.4	99.3	722.5	1734.6	29311.1
avg. pixel NTL (radiance)	1.0	0.002	0.02	0.07	0.2	0.4	1.5	20.3
LIDCs								
NTL growth	7.4	-54.2	-16.7	-3.7	6.6	18.1	32.0	78.4
GDP growth	3.7	-17.7	-1.4	2.2	4.5	6.4	8.3	13.9
avg. # of obs per year	101.2	43.1	59.9	71.6	103.9	126.1	144.9	159.3
area (thousand sq. km)	360.8	1.4	21.9	59.4	197.2	507.2	928.7	1900.0
avg. pixel NTL (radiance)	0.04	0.0003	0.002	0.005	0.01	0.04	0.09	0.3
Total								
NTL growth	6.0	-61.1	-15.5	-4.3	4.5	15.1	29.5	85.2
GDP growth	3.3	-14.5	-1.7	1.4	3.6	5.7	7.9	15.9
avg. # of obs per year	102.4	39.7	56.2	71.3	98.3	129.6	159.1	190.1
area (thousand sq. km)	811.7	0.07	1.7	23.2	135.6	535.6	1476.5	16864.8
avg. pixel NTL (radiance)	1.0	0.0008	0.005	0.02	0.1	0.5	1.7	18.8

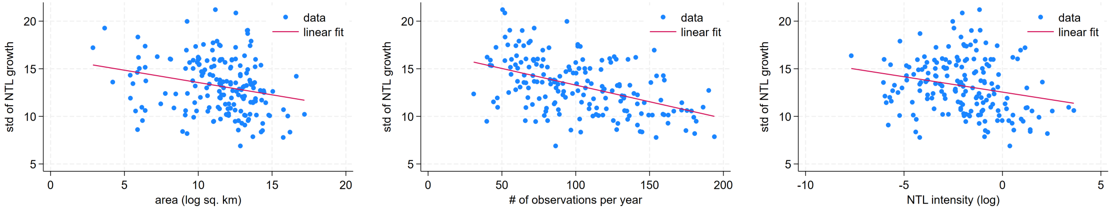
181 countries, 1993–2023. NTL growth is an order of magnitude more volatile than GDP growth.

Summary Statistics: VIIRS Quarterly Dataset

	Mean	p1	p10	p25	p50	p75	p90	p99
1st quarter								
NTL growth (yoy)	3.8	-86.2	-24.1	-7.2	3.8	14.7	31.5	107.0
GDP growth (yoy)	2.7	-7.1	-1.1	1.0	2.7	4.6	6.5	13.3
avg. # of obs	27.9	7.4	13.6	24.1	27.4	33.6	38.0	43.6
avg. pixel NTL	0.9	0.02	0.09	0.2	0.3	0.5	1.3	12.9
2nd quarter								
NTL growth (yoy)	9.6	-200.5	-22.4	-4.3	5.6	19.0	45.0	272.6
GDP growth (yoy)	2.8	-22.9	-2.6	1.2	3.2	5.4	8.1	20.8
avg. # of obs	18.3	0.1	2.9	7.3	13.4	29.4	43.1	48.7
avg. pixel NTL	0.8	0.001	0.04	0.09	0.2	0.3	0.8	19.5
3rd quarter								
NTL growth (yoy)	4.9	-150.0	-31.7	-9.8	4.3	18.5	44.6	157.7
GDP growth (yoy)	2.8	-10.9	-1.9	1.2	3.0	5.0	7.1	15.4
avg. # of obs	22.1	0.7	4.9	10.7	17.3	33.7	45.4	53.9
avg. pixel NTL	0.7	0.005	0.04	0.09	0.2	0.3	0.8	18.8
4th quarter								
NTL growth (yoy)	6.5	-76.4	-18.3	-4.7	4.3	14.7	33.3	132.9
GDP growth (yoy)	2.8	-9.0	-1.5	1.1	2.8	4.8	6.7	12.6
avg. # of obs	29.6	6.8	15.8	25.7	29.2	35.6	42.9	48.0
avg. pixel NTL	0.8	0.02	0.09	0.2	0.3	0.5	1.2	11.8
Total								
NTL growth (yoy)	6.2	-120.8	-24.1	-6.5	4.5	16.4	38.2	158.8
GDP growth (yoy)	2.8	-14.6	-1.6	1.1	2.9	4.9	7.0	16.8
avg. # of obs	24.6	0.5	7.1	13.6	26.1	33.9	41.9	51.5
avg. pixel NTL	0.8	0.008	0.07	0.1	0.2	0.4	0.9	18.8

107 countries, 2013Q2–2025Q2. Strong seasonal patterns in NTL observations.

Evidence: Heteroskedasticity of NTL Growth



(a) vs. Area (b) vs. N^{obs} (c) vs. NTL intensity

Figure: Standard deviation of NTL growth vs. country characteristics (SVNL)

Larger, more frequently observed, and brighter countries have less volatile NTL growth.

Heteroskedasticity: Regression Evidence

Table: Volatility vs. Country Characteristics

	SVNL annual		VIIRS quarterly	
	(1)	(2)	(3)	(4)
	std(z)	std(y)	std(z)	std(y)
area (log sq. km)	-0.381***	-0.066	-0.108	-0.231***
	(0.073)	(0.041)	(0.165)	(0.045)
avg. # obs per year	-0.031***	-0.001	-0.155***	0.007
	(0.005)	(0.003)	(0.025)	(0.007)
NTL intensity (log)	-0.472***	-0.060	-1.901***	-0.387***
	(0.095)	(0.054)	(0.300)	(0.082)
constant	19.750***	4.278***	19.650***	5.374***
	(0.887)	(0.503)	(1.783)	(0.489)
Obs	181	181	373	373
Adjusted R^2	0.338	0.000	0.256	0.068

Country characteristics explain NTL growth volatility (high R^2) but not GDP growth volatility.

Naive Structural Relationship: SVNL Annual

	NTL growth				
	(1)	(2)	(3)	(4)	(5)
	full sample	full sample	AEs	EMs	LIDCs
GDP growth	0.482***	0.490***	0.457***	0.467***	0.404***
	(0.0478)	(0.0477)	(0.125)	(0.0631)	(0.0914)
constant	3.963***	3.935***	1.987***	4.141***	5.431***
	(0.248)	(0.235)	(0.524)	(0.334)	(0.516)
year fixed effects	-	Yes	-	-	-
Obs	4818	4818	961	2433	1424
Adjusted R^2	0.0204	0.193	0.0127	0.0216	0.0129

Coefficients are biased downward due to measurement errors. Serve as lower bound for the elasticity.

Naive Structural Relationship: VIIRS Quarterly

	NTL growth				
	(1)	(2)	(3)	(4)	(5)
	full sample	full sample	AEs	EMs	LIDCs
GDP growth	0.194***	0.224***	0.328**	0.208**	-0.394*
	(0.073)	(0.077)	(0.150)	(0.093)	(0.202)
constant	4.914***	4.828***	2.080***	5.726***	9.143***
	(0.340)	(0.335)	(0.581)	(0.441)	(1.168)
year and quarter fixed effects	-	Yes	-	-	-
Obs	3964	3964	1144	2289	531
Adjusted R^2	0.002	0.127	0.003	0.002	0.005

Quarterly coefficients are even smaller — greater noise in quarterly GDP biases them further.

Predictive Relationship

SVNL Annual

	GDP growth				
	(1) full sample	(2) full sample	(3) AEs	(4) EMs	(5) LIDCs
NTL growth	0.0429***	0.0441***	0.0301***	0.0471***	0.0335***
	(0.00425)	(0.00429)	(0.00823)	(0.00637)	(0.00759)
constant	3.118***	3.111***	2.464***	3.023***	3.773***
	(0.0614)	(0.0569)	(0.110)	(0.0905)	(0.117)
year fixed effects	-	Yes	-	-	-
Obs	4818	4818	961	2433	1424
Adjusted R^2	0.0204	0.184	0.0127	0.0216	0.0129

VIIRS Quarterly

	GDP growth				
	(1) full sample	(2) full sample	(3) AEs	(4) EMs	(5) LIDCs
NTL growth	0.0092***	0.0095***	0.0128**	0.0105**	-0.0181*
	(0.0035)	(0.0033)	(0.0058)	(0.0047)	(0.0093)
constant	2.814***	2.813***	2.205***	2.812***	4.283***
	(0.061)	(0.055)	(0.095)	(0.084)	(0.187)
year and quarter fixed effects	-	Yes	-	-	-
Obs	3964	3964	1144	2289	531
Adjusted R^2	0.0015	0.221	0.0033	0.0017	0.0052

Predictive coefficients can be directly used for GDP prediction where official data is missing.

Baseline Results: Estimated Parameters

Table: Baseline GMM Estimates

	β	σ_{y^*}	σ_{ε^y}	$\log \sigma_z^2$	σ_{η^z}
SVNL annual					
Point Estimate	1.52***	4.19**	3.12*	-9.24***	13.24***
	(0.41)	(1.95)	(1.84)	(3.16)	(0.47)
t-statistic	3.73	2.15	1.69	-2.93	27.89
VIIRS quarterly					
Point Estimate	2.25***	2.90	3.71	-6.01	17.11***
	(0.54)	(1.83)	(3.42)	(16.19)	(1.74)
t-statistic	4.19	1.59	1.09	-0.37	9.84

- Elasticity: $\hat{\beta} = 1.52$ (annual), $\hat{\beta} = 2.25$ (quarterly)
- 95% CI: $[0.72, 2.31]$ (annual), $[1.20, 3.31]$ (quarterly)
- Not statistically different from each other

Comparison with Previous Estimates

Study	Data	Elasticity
Henderson, Storeygard, and Weil (2012)	DMSP-OLS	1.00
Hu and Yao (2022)	DMSP-OLS	1.30
This paper (annual)	**SVNL**	**1.52**
This paper (quarterly)	**VIIRS**	**2.25**

- Our estimates are higher, but the differences are statistically small given confidence intervals
- Higher VIIRS estimates may reflect improved low-light sensitivity

Signal-to-Total Variance Ratio of GDP

$$\text{Sig}_y = \frac{\sigma^2_{y^*}}{\sigma^2_{y^*} + \sigma^2_{\varepsilon y}}$$

	AEs	EMs	LIDCs
Annual (SVNL)	96%	62%	56%
Quarterly (VIIRS)	58%	39%	25%

- GDP accuracy decreases with development level
- Quarterly GDP is much noisier than annual GDP
- Henderson, Storeygard, and Weil (2012): assumed 90% for AEs, derived 59% for "bad data" countries — close to our annual estimates
- Our results follow naturally without assumptions on GDP errors

Optimal Weights for Light-Adjusted GDP

Using the estimated parameters and the predictive coefficient γ:

	Annual	Quarterly
Optimal weight λ^*	0.38	0.63

- For a country with median characteristics
- Higher quarterly weight reflects noisier official GDP at higher frequency
- NTL plays a meaningful role in improving GDP estimates

Elasticity by Country Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Income status		GDP per capita		Agriculture share		Industry share		Services share	
	AEs	EMDEs	Below	Above	Below	Above	Below	Above	Below	Above
SVNL										
β	1.99**	1.57***	1.45***	1.17	1.00***	1.51***	1.64***	1.81***	1.27***	0.70
	(0.91)	(0.37)	(0.37)	(0.73)	(0.38)	(0.35)	(0.39)	(0.35)	(0.34)	(0.62)
t-statistic	2.20	4.22	3.96	1.60	2.66	4.29	4.23	5.11	3.72	1.12
VIIRS										
β	4.55***	2.85***	2.26***	3.23**	2.48**	1.12	5.08***	2.49***	1.71**	-0.54
	(1.62)	(0.63)	(0.75)	(1.27)	(1.17)	(1.55)	(1.13)	(0.71)	(0.76)	(2.03)
t-statistic	2.81	4.53	3.00	2.55	2.12	0.72	4.49	3.50	2.26	-0.26

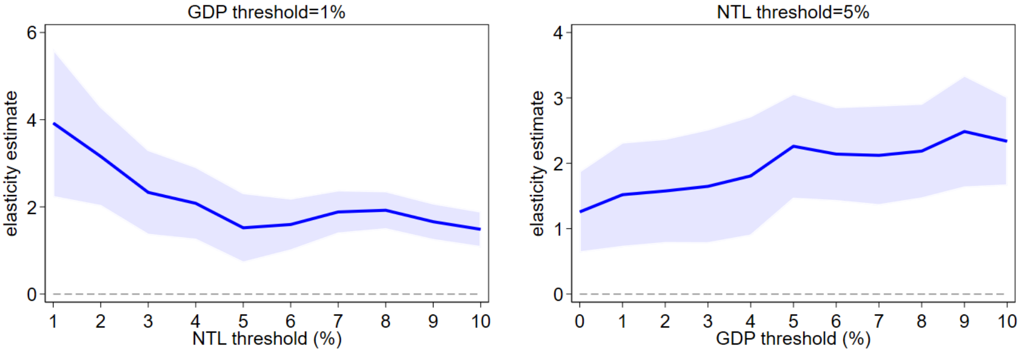
Elasticity is relatively stable across groups. Differences are modest given standard errors.

Elasticity Heterogeneity: Key Takeaways

- **Income status**: AEs slightly higher than EMDEs, but imprecisely estimated
- **GDP per capita**: Above-median not statistically significant (annual)
- **Agriculture share**: Even above-average agriculture countries show significant elasticity (annual)
- **Service share**: Above-average service share weakens the relationship
- **Quarterly**: Generally higher elasticity; weakens for high-agriculture and high-service economies

Overall: NTL–GDP relationship is robust across most economies at both frequencies

Robustness: Winsorization (SVNL Annual)

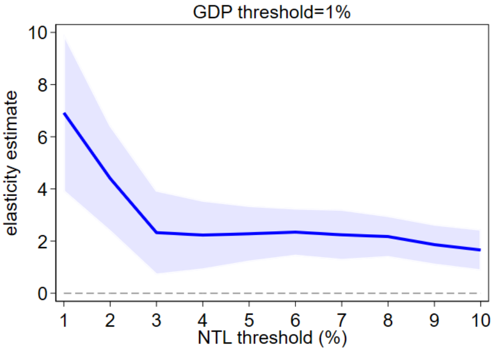


(a) Varying NTL threshold (GDP threshold = 1%) (b) Varying GDP threshold (NTL threshold = 5%)

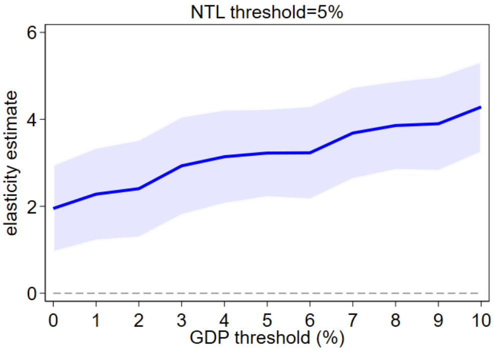
Figure: SVNL Annual: Elasticity under various winsorization

Elasticity stabilizes at NTL threshold $\geq 5\%$. Robust to GDP threshold variation.

Robustness: Winsorization (VIIRS Quarterly)



(a) Varying NTL threshold (GDP threshold = 1%)



(b) Varying GDP threshold (NTL threshold = 5%)

Figure: VIIRS Quarterly: Elasticity under various winsorization

Same pattern: winsorization of NTL tails is essential for stable estimation.

Summary

1. **New framework**: identification through heteroskedasticity of NTL measurement errors
   - Grounded in the satellite data-generating process
   - No assumptions on GDP quality needed
2. **Elasticity estimates**: 1.52 (annual), 2.25 (quarterly)
   - Stable across country groups and sectors
   - Robust to winsorization choices
3. **GDP signal ratios**: 96% (AEs) $\rightarrow$ 56% (LIDCs) annually
   - NTL can substantially improve GDP statistics for developing countries
   - Quarterly GDP is much noisier; optimal weight on NTL is higher
4. **General applicability**: framework extends to other satellite data with heteroskedastic errors

Thank you!

References I

Chen, X., and W. D. Nordhaus (2011): "Using luminosity data as a proxy for economic statistics," *Proceedings of the National Academy of Sciences*, 108(21), 8589–8594.

Chen, X., Z. Wang, F. Zhang, G. Shen, and Q. Chen (2024): "A global annual simulated VIIRS nighttime light dataset from 1992 to 2023," *Scientific Data*, 11(1), 1380.

Civelli, A., A. Gaduh, and A. S. Yousuf (2024): "Measuring Economic Growth with a Fully Identified Three-Signal Model," *Review of Economics and Statistics*, pp. 1–30.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012): "Measuring Economic Growth from Outer Space," *American Economic Review*, 102(2), 994–1028.

Hu, Y., and J. Yao (2022): "Illuminating economic growth," *Journal of Econometrics*, 228(2), 359–378.

Backup Slides

Derivation: Variance of NTL Growth Error

Temporal averaging: $\varepsilon_{p,t} = \frac{1}{N_i^{obs}} \sum_d \varepsilon_{p,d} \Rightarrow \text{var}(\varepsilon_{p,t}) = \sigma_z^2 / N_i^{obs}$

Spatial aggregation: $\text{var}(\varepsilon_{i,t}) = N_i \cdot \sigma_z^2 / N_i^{obs}$

Growth rate (Taylor expansion of log): $\log(Z_{i,t}) \approx \log(Z_{i,t}^*) + \varepsilon_{i,t} / Z_{i,t}^*$

Result:

$$\begin{aligned} \text{var}(\varepsilon_{i,t}^z) &\approx \frac{\text{var}(\varepsilon_{i,t})}{(Z_{i,t}^*)^2} + \frac{\text{var}(\varepsilon_{i,t-1})}{(Z_{i,t-1}^*)^2} \\ &\approx \frac{2N_i\sigma_z^2 / N_i^{obs}}{(\mathrm{E}z_i N_i)^2} = \frac{2\sigma_z^2}{N_i \cdot N_i^{obs} (\mathrm{E}z_i)^2} \end{aligned}$$

Derivation: Optimal Weight

MSE of light-adjusted GDP:

$$E(\bar{y} - y^*)^2 = (1-\lambda)^2 \operatorname{var}(\varepsilon^y) + \lambda^2 \left[(\gamma\beta - 1)^2 \operatorname{var}(y^*) + \gamma^2 \operatorname{var}(\mu^z) \right]$$

FOC yields:

$$\lambda^* = \frac{\operatorname{var}(\varepsilon^y)}{\operatorname{var}(\varepsilon^y) + (\gamma\beta - 1)^2 \operatorname{var}(y^*) + \gamma^2 \operatorname{var}(\mu^z)}$$

All components are identified:

$$\begin{aligned} \operatorname{var}(y^*) &= \operatorname{cov}(y, z)/\beta \\ \operatorname{var}(\varepsilon^y) &= \operatorname{var}(y) - \operatorname{var}(y^*) \\ \operatorname{var}(\mu^z) &= \operatorname{var}(z) - \beta^2 \operatorname{var}(y^*) \end{aligned}$$