Essays on Oil, Attention Disorders and E-commerce Thresholds

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Lucas Chaves Gurgel

August 2026

© by Lucas Chaves Gurgel, 2026
All Rights Reserved.

To Luciane Gurgel

Acknowledgements

I would like to thank my wife, Luciane, and my children, José Miguel, Catarina, Ana Clara, Heloísa, Thomas, and most recently Maria Sophia, for their unwavering support and inspiration throughout this journey. This work would not have been possible without your love and encouragement.

I am deeply grateful for the patience, wisdom, and guidance provided by my dissertation committee: Celeste Carruthers, Don Bruce, Luiz Lima, Andrew Hanson, and Robert Kelchen. Your mentorship has shaped me as a scholar and economist.

My sincere thanks extend to the University of Tennessee's Boyd Center for Business and Economic Research and the Department of Economics for their generous financial support and for fostering such a welcoming and intellectually stimulating environment. I am fortunate to have been surrounded by wonderful professors, colleagues, and friends: Georg Schaur, Scott Holladay, Larry Kessler, Mike Kofoed, Christian Vossler, Kira Rasmussen, Sherri Pinkston, Beth Weissmueller, Ed Taylor, Donna Bueckman, Ken Baker, Mohammed Mohsin, James Lake, and Matt Van Essen.

To my cohort and fellow graduate students who shared this road with me — Kaden Grace, Connor Neff, Jonathan Paul Lynn, Noehmi Torres, Deng Pen, Trenton Marable, and Thomas Vierville — thank you for the camaraderie, the conversations, and the countless moments that made this experience truly memorable.

I also owe a great deal to the faculty and colleagues at the Federal University of Ceará who first encouraged me to pursue this path: Elano Arruda, Ivan Castelar, and Dércio de Assis. Your belief in my potential planted the seed for everything that

followed. I am equally grateful to the Federal University of Ceará for its financial support during my time abroad.

Finally, I wish to express my deepest gratitude to my parents, Ana and Tércio, and my parents-in-law, Lúcia and Nelson, for their unconditional love, emotional support, and material sacrifices throughout this long journey. Your dedication has been my foundation.

Abstract

This dissertation comprises three empirical essays examining how technological, regulatory, and commodity market forces shape economic outcomes across macroeconomic and microeconomic dimensions.

The first essay investigates the relationship between oil price fluctuations and macroeconomic tail risk. Using a quantile regression framework, I model the conditional distribution of GDP growth incorporating oil price movements, financial conditions, and current output. The model improves forecast accuracy of the lowest quantiles of future GDP growth by 12%. Counterfactual analysis reveals that high oil price fluctuations can increase the probability of a recession by 60% and future GDP growth variance by 20%, highlighting oil's role in generating economic uncertainty beyond its effects on average growth.

The second essay examines the causal impact of high-speed internet connectivity on ADHD outcomes in the United States. Using county-level data from 2008 to 2023 and exploiting variation in broadband adoption timing, I employ two-way fixed effects models to estimate the effect of internet penetration on demand for ADHD treatment medications. In the main specification, which uses the Federal Communications Commission (FCC) internet penetration scale (0–5), each increase in broadband adoption tier is associated with a 0.85 percent increase in total ADHD medication demand, with amphetamine-based medications showing the highest increase at 1.65 percent. Effects are more pronounced in smaller counties. Evidence from the American Time Use Survey suggests that youth, particularly boys, increase computer

screen time as broadband expands, pointing to a behavioral mechanism linking digital infrastructure to ADHD diagnoses. These findings carry implications for public health policy surrounding the broader consequences of internet expansion.

The third essay analyzes the economic consequences of sales tax enforcement following the *South Dakota v. Wayfair* (2018) Supreme Court decision, which enabled states to impose tax obligations on remote sellers without physical presence. Using a state-level panel from 2013 to 2023, I exploit cross-state variation in economic nexus threshold tightness relative to market size. I find modest but statistically significant negative effects on the number of firms, establishments, and employment, concentrated among small businesses. These results suggest that less binding thresholds create competitive disadvantages for physical retailers and underscore that threshold exposure, rather than threshold level alone, is the policy-relevant parameter for achieving competitive equity between online and brick-and-mortar commerce*.

*Artificial intelligence tools were employed for text revision, grammatical correction, and linguistic refinement to enhance clarity and fluency. All conclusions, opinions, and any remaining errors are my own.

Table of Contents

1 Introduction **1**

2 Surfing the Black Wave: Assessing Oil Prices and Macroeconomic Tail Risk **7**

2.1 Oil Prices and Macroeconomic Tail Risk 9

2.2 Empirical Framework 11

2.2.1 Data 11

2.2.2 Quantile Estimation Approach 12

2.2.3 Quantile Estimation with Mixed Frequency Data 14

2.2.4 Out-of-sample Evaluation 14

2.2.5 Conditional Future GDP Growth 16

2.2.6 Oil and GDP Downside Risk 16

2.3 Results and Discussion 17

2.3.1 In-sample Quantile Estimates 17

2.3.2 Out of Sample Performance 19

2.3.3 Density Forecast 24

2.3.4 Vulnerability Measures 27

2.3.5 The COVID-19 Pandemic and Structural Break 29

2.4 Counterfactual Analysis - CoVaR 31

2.5 Evidence on the Mechanism 37

2.6 Discussion 42

3 Connected and Distracted: The Impact of High-Speed Internet on ADHD Outcomes in the U.S. 47

3.1 High-Speed Internet Evolution and Attention-Deficit Disorders 49

3.2 Empirical Framework 53

3.2.1 Data 53

3.2.2 Background and National Trends 55

3.2.3 Econometric Approach 62

3.3 Results 64

3.4 Alternative Specifications 69

3.5 Suggestive Evidence on Mechanisms 75

3.6 Discussion 81

4 Sales Tax Enforcement and Business Activity: Evidence from Economic Nexus Laws 92

4.1 E-commerce Taxation and Business Activity 94

4.2 Legal Background & Literature 99

4.3 Conceptual Framework 103

4.4 Empirical Framework 105

4.4.1 Data Sources 105

4.4.2 Descriptive Evidence 107

4.4.3 Econometric Approach 109

4.5 Results 110

4.5.1 Heterogeneity by Size and Age 111

4.5.2 Heterogeneity by Industry Type 113

4.6 Alternative Specifications 116

4.6.1 Fixed Population Denominator 116

4.6.2 Market Size Alternative Measure 116

4.6.3 Gradual Addition of Controls 117

4.6.4 Adoption Timing Check 121

4.7 Discussion 121

Vita **126**

List of Tables

2.1 Data Summary 13
2.2 Specification Summary 18
2.3 Unconditional GDP Growth 18
2.4 Quantile Regression Results - One Quarter Ahead 20
2.5 Regression Quantile Results - One Year Ahead 21
2.6 Pseudo R2 OOS Results - One Quarter Ahead 22
2.7 Pseudo R2 Out of Sample Results - One Year Ahead 23
2.8 Berkowitz test statistics 25
2.9 Average of Vulnerability Measures - Out-of-Sample Analysis 28
2.10 Average of Vulnerability Measures - In-Sample Analysis 28
2.11 Regression Quantile Results - One Quarter Ahead (Excl. COVID 2020Q1–2020Q4) 30
2.12 CoVaR – Moments of the Density Forecast 35
2.13 CoVaR – Tail Risk Measures 36
2.14 Oil Price Coefficient on GDP Forecast Dispersion Across Quantiles . 40
2.15 Oil Price Coefficient on Consumer Confidence Growth Across Quantiles 40
2.16 Oil Price Coefficient on Total Capacity Utilization Across Quantiles . 44

3.1 Two-Way Fixed Effects: High-Speed Internet and ADHD Treatment Demand 66
3.2 Two-Way Fixed Effects: High-Speed Internet and IEP Students . . . 66
3.3 Alternative Time-Trend Controls 70

3.4 High-Speed Internet on ADHD Drug Demand by Different Thresholds 76
3.5 High-Speed Internet and ADHD Medications by Drug Type 77
3.6 TWFE – High-Speed Internet and Computer Time Use 84
3.7 Time Period Sub Groups – Heterogeneous Treatment Effects 85
3.8 Low-Speed Internet Penetration and ADHD Treatment Drug Demand 86

4.1 Impact of Economic Nexus Threshold Exposure 112
4.2 Impact of Threshold Exposure by Firm Size 112
4.3 Impact of Threshold Exposure by Firm Age 114
4.4 Estimation Results by Industrial Sector 115
4.5 Alternative Treatment: Threshold Scaled by Fixed 2017 Population . 118
4.6 Alternative Treatment: Threshold Scaled by Goods-Producing Consumption GDP 118
4.7 Baseline Estimation Results with Gradual Inclusion of Controls . . . 120
4.8 Frozen Treatment Results (Log-Lin Specification, Staggered Adoption Check) 125

List of Figures

2.1 Oil Prices - From 1974 until 2024 13
2.2 Unconditional GDP Growth Distribution 25
2.3 Conditional GDP Growth Distribution, One Quarter Ahead: 2008Q4 26
2.4 Conditional GDP Growth Distribution, One Quarter Ahead: 2024Q1 26
2.5 Growth-at-Risk measure - One quarter ahead 28
2.6 $\beta_{Oil|\tau=0.05}$ - Oil Estimate and Significance over time 28
2.7 $\beta_{Oil|\tau=0.05}$ - Oil Estimate and Significance over time (Excl. COVID 2020Q1–2020Q4) 30
2.8 Impact of Higher Oil Price Variation on Conditional GDP Growth . . 33

3.1 Evolution of High-Speed Internet Connectivity in the United States (2008–2023) 56
3.2 Geographic Distribution of High-Speed Internet Access by Year . . . 57
3.3 ADHD Medication Distribution Trends by Drug Type (2008–2023) . 60
3.4 Evolution of IEP Ratio in the United States (2009–2019) 61
3.5 Number of Treated and Untreated Counties by Year 67
3.6 Event-Study Estimates – DID Imputation (Borusyak et al., 2024) . . 68
3.7 Counties Treatment Status per Tier-Threshold 72
3.8 Heterogeneity Analysis by County Population Size 74
3.9 High-Speed Internet Adoption and Computer Time by Age Group . . 76

4.1 Adoption of Economic Nexus Thresholds Over Time 96

4.2 Economic Nexus Threshold Exposure Across U.S. States (2023) . . . 97
4.3 Competing Effects of Economic Nexus Thresholds on Business Activity 104
4.4 Overall Business Activity Trends in the United States, 2013–2023 . . 108
4.5 Threshold per $1 Million of Goods-Producing Consumption GDP Across States (2023) 119

Chapter 1

Introduction

This dissertation consists of three distinct research projects covering applied macroeconomics, labor economics, and public economics that collectively demonstrate the breadth and methodological diversity of modern empirical economics. Each chapter employs rigorous econometric techniques to address pressing questions within its respective domain, showcasing the versatility of economic analysis in understanding complex real-world phenomena across multiple scales and contexts.

In the second chapter, I characterize how oil price information affects the conditional distribution of future GDP growth, with particular focus on extreme negative outcomes during recession periods. There is a long tradition of economic research evaluating the impact of oil price fluctuations on macroeconomic performance, with more recent literature identifying a diminishing relevance of oil price shocks to the general performance of the U.S. economy (Blanchard & Gali, 2007; Hamilton, 1996; Kilian & Vigfusson, 2017). Yet this literature has largely focused on average effects, leaving open the question of whether oil price information matters more in the tails of the growth distribution—precisely where policy concern is greatest.

I contribute to a growing literature on growth-at-risk (Adrian et al., 2019; Loria et al., 2023) by incorporating a natural resource component into the analysis of GDP tail risk. Using a quantile regression framework (Koenker & Bassett Jr, 1978), I demonstrate that oil price fluctuations carry meaningful information for forecasting the lowest quantiles of future GDP growth, while having limited relevance for median and upper quantiles representing normal economic conditions. A counterfactual exercise extending the ΔCoVaR framework of Adrian and Brunnermeier (2016) to the macroeconomic setting further illustrates how extreme oil price variations can substantially amplify economic uncertainty, motivating greater attention to asymmetric distributional effects when evaluating commodity price shocks.

In the third chapter, I examine the unintended consequences of broadband expansion on childhood attention disorders and special education needs. The proliferation of high-speed internet has reshaped how individuals allocate time and conduct daily activities, generating a complex landscape of opportunities and

challenges for human development. While research demonstrates that internet access can improve educational outcomes and human capital formation (Dettling et al., 2018; Grimes & Townsend, 2018), emerging evidence reveals concerning negative impacts on health and cognitive development (Arenas-Arroyo et al., 2025; Colombo & Failache, 2023). Early medical studies have begun to associate increased screen exposure and hyperstimulation with difficulties in concentration and learning disabilities such as Attention Deficit Hyperactivity Disorder (ADHD), with children who have greater exposure to television and video games showing higher rates of attention problems (Swing et al., 2010).

Fletcher (2014) demonstrates that childhood ADHD significantly impairs adult labor market outcomes, with affected individuals experiencing reduced employment rates and lower wages throughout their careers. Leveraging the staggered rollout of high-speed internet adoption across U.S. counties as a quasi-natural experiment, I investigate how increased connectivity affects ADHD medication demand and special education enrollment. The analysis combines data from the Federal Communications Commission, DEA pharmaceutical distribution records, and educational administrative datasets, using modern causal inference methods designed for staggered treatment adoption. This research speaks to broader discussions about the unintended fiscal and developmental consequences of federal broadband infrastructure programs such as the Connect America Fund, which has directed over $10 billion toward expanding digital access in underserved communities.

In the fourth chapter, co-authored with Don Bruce, we investigate whether variation in economic nexus thresholds for e-commerce taxation generates distinct economic consequences following the landmark *South Dakota v. Wayfair* (2018) decision. Prior to *Wayfair*, the physical presence rule established in *Quill Corp. v. North Dakota* (1992) created a long-standing tax advantage for remote sellers, prompting scholars to call for reform and raising concerns about the erosion of state tax bases (Agrawal, 2021; Bruce & Fox, 2000). The *Wayfair* decision has shifted the landscape toward destination-based taxation, and nearly every U.S. state

subsequently adopted economic nexus thresholds—most converging toward $100,000 in annual sales.

The selection of these thresholds appears somewhat arbitrary, yet their economic consequences for business activity and interstate commerce remain poorly understood. We construct a novel measure of threshold exposure—defined as the threshold amount normalized by state population—to capture the degree to which a given threshold is binding relative to local market size. Exploiting cross-state variation in threshold exposure using a state-level panel from 2013 to 2023, we employ two-way fixed effects models to evaluate the impact of threshold design on firm activity and employment. This research addresses whether the seemingly arbitrary selection of economic nexus thresholds has measurable effects on local economic activity and competitive equity between online and brick-and-mortar retailers.

Each chapter stands as an independent contribution to its respective field. Jointly, these studies offer a broad perspective on how economic shocks, policy design choices, and technological interventions can generate asymmetric distributional effects, contributing to ongoing discussions about risk management, optimal policy coordination, and the comprehensive evaluation of public programs. Further details on each study are provided in the chapters that follow.

References

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, *109*(4), 1263–1289.

Adrian, T., & Brunnermeier, M. K. (2016). Covar. *The American Economic Review*, *106*(7), 1705.

Agrawal, D. R. (2021). The internet as a tax haven? *American Economic Journal: Economic Policy*, *13*(4), 1–35.

Arenas-Arroyo, E., Fernandez-Kranz, D., & Nollenberger, N. (2025). High speed internet and the widening gender gap in adolescent mental health: Evidence from spanish hospital records [Forthcoming]. *Journal of Health Economics*. https://doi.org/10.1016/j.jhealeco.2025.103014

Blanchard, O. J., & Gali, J. (2007). The macroeconomic effects of oil shocks: Why are the 2000s so different from the 1970s?

Bruce, D., & Fox, W. F. (2000). E-commerce in the context of declining state sales tax bases. *National Tax Journal*, *53*(4), 1373–1388.

Colombo, K., & Failache, E. (2023, November). *Exposure to high-speed internet and early childhood development: Evidence from a countrywide program* (Working Paper No. 2023-01) (First version: January 19, 2023. This version: November 1, 2023). GEAR - Graduate Program in Applied Economic Research, Universitat Autònoma de Barcelona.

Dettling, L. J., Goodman, S., & Smith, J. (2018). Every little bit counts: The impact of high-speed internet on the transition to college. *The Review of Economics and Statistics*, *100*(2), 260–273. https://doi.org/10.1162/rest_a_00712

Fletcher, J. M. (2014). The effects of childhood adhd on adult labor market outcomes. *Health Economics*, *23*(2), 159–181. https://doi.org/10.1002/hec.2907

Grimes, A., & Townsend, W. (2018). Effects of (ultra-fast) fibre broadband on student achievement. *Information Economics and Policy*, *44*, 8–15. https://doi.org/10.1016/j.infoecopol.2018.06.001

Hamilton, J. D. (1996). This is what happened to the oil price-macroeconomy relationship. *Journal of monetary economics*, *38*(2), 215–220.

Kilian, L., & Vigfusson, R. J. (2017). The role of oil price shocks in causing us recessions. *Journal of Money, Credit and Banking*, *49*(8), 1747–1776.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 33–50.

Loria, F., Matthes, C., & Zhang, D. (2023). Assessing macroeconomic tail risk. *Available at SSRN 4002665*.

Swing, E. L., Gentile, D. A., Anderson, C. A., & Walsh, D. A. (2010). Television and video game exposure and the development of attention problems. *Pediatrics*, *126*(2), 214–221. https://doi.org/10.1542/peds.2009-1508

Chapter 2

Surfing the Black Wave: Assessing Oil Prices and Macroeconomic Tail Risk

Abstract

I study the impact of oil price fluctuations on the conditional distribution of U.S. GDP growth using a quantile regression framework. I propose a model incorporating oil price fluctuations, financial conditions, and current economic activity, and show that it improves out-of-sample forecast accuracy for the lower tail of future GDP growth (Growth-at-Risk) by approximately 12% at a one-quarter-ahead horizon; the effect dissipates at longer horizons, with no significant impact found one year ahead. Oil price fluctuations primarily affect downside risk: they thicken the left tail of the conditional GDP growth distribution without shifting the center or upper tail. A counterfactual exercise extending ΔCoVaR to a macroeconomic setting shows that large oil price fluctuations increase the probability of recession by up to 60% and raise the variance of future GDP growth by approximately 20%. These results indicate that oil price fluctuations act as a short-run source of macroeconomic tail risk, asymmetrically amplifying the likelihood of adverse growth outcomes.

2.1 Oil Prices and Macroeconomic Tail Risk

Oil price fluctuations have long been regarded as a driver of macroeconomic instability. The 1973 OPEC embargo, the 1979 supply shock, and the 2008 price surge each preceded sharp contractions in U.S. output, establishing oil as a recurring source of macroeconomic disruption. A large literature evaluates the macroeconomic effects of oil price shocks, with recent work finding a diminishing sensitivity of U.S. output to oil price fluctuations (Hamilton, 1996; Blanchard and Gali, 2007; Kilian and Vigfusson, 2017). Most of this literature, however, focuses on the mean or median response of output to oil shocks. In this paper, I use quantile regression (Koenker & Bassett Jr, 1978) to assess the impact of oil price fluctuations on the conditional distribution of future GDP growth, with particular interest in the left tail and distributional measures of uncertainty such as variance.

A growing body of research investigates variables that help explain macroeconomic tail risk — the probability of severe economic downturns concentrated in the left tail of the output growth distribution. Drawing from finance, the growth-at-risk (GaR) framework applies value-at-risk logic to GDP, defining tail risk as the value of economic growth at its 5th percentile (Sui et al., 2024; Aikman et al., 2019; IMF, 2017). Adrian et al. (2019) characterize the GDP growth distribution using quantile regression and the NFCI, finding that tighter financial conditions are strongly associated with a higher probability of negative GDP growth. Loria et al. (2023) combine quantile regression with local projections (Jordà, 2005) to estimate the effects of monetary policy, financial conditions, uncertainty, and oil price shocks on the predictive distribution of Industrial Production growth, finding that the 10th percentile is three to six times more sensitive than the median.

In the wake of the Global Financial Crisis (GFC), the measurement of systemic risk has become a central concern in macroeconomic research. Adrian and Brunnermeier (2016) propose ΔCoVaR, a reduced-form measure of systemic risk defined as the change in the value-at-risk of the financial system conditional on a particular

institution being in distress (returns below the $\tau = 0.10$ quantile) relative to its median state. Estimated using weekly equity returns from 1971 to 2013, they find that institutional characteristics such as leverage, size, and maturity mismatch significantly predict ΔCoVaR in the cross section. They also construct a forward-looking version of the measure that is countercyclical — systemic risk builds up during tranquil periods and materializes in crises — and show that its 2006 estimates predicted more than one-third of the realized ΔCoVaR during the 2007–2009 financial crisis. Giglio et al. (2016) extend this agenda by evaluating 19 systemic risk measures across the US and Europe, finding that increases in systemic risk skew the distribution of future macroeconomic shocks asymmetrically — shifting the left tail downward without a corresponding effect on the median. Aggregating individual measures into a systemic risk index via partial quantile regression, they show that a one standard deviation increase in systemic risk shifts the 20th percentile of industrial production growth by more than 50%, while predictive power for the central tendency remains weak.

I find that oil price fluctuations carry no predictive content for the median or upper quantiles of GDP growth, but improve forecast accuracy at the lower quantiles — outperforming the benchmark by 12.77% for the Growth-at-Risk prediction ($\tau < 0.05$). Using a t-skew distribution (Azzalini & Capitanio, 2003), I estimate the GDP growth distribution conditional on oil price fluctuations, financial conditions, and current economic activity, which allows me to assess how the probability of negative growth responds to oil price fluctuations across time. I also estimate distributional characteristics beyond tail effects — variance and interquartile range — quantifying aggregate uncertainty following oil price shocks. Extending Adrian and Brunnermeier (2016)'s ΔCoVaR to a macroeconomic setting, I find that extreme oil price volatility increases the probability of recession from 10% to 16% and raises the variance of the GDP growth distribution by 20%.

This paper contributes to the existing literature in two ways. First, it shifts the lens through which oil price fluctuations are evaluated — away from their effect on mean or median output and toward their impact on tail risk and aggregate

uncertainty. The evidence suggests that the relevance of oil price fluctuations for macroeconomic outcomes has been understated. In particular, incorporating oil price fluctuations meaningfully improves the prediction of worst-case growth outcomes, outperforming standard benchmarks. Second, I extend Adrian and Brunnermeier (2016)'s ΔCoVaR to a macroeconomic setting, providing a tractable framework for quantifying how extreme oil price volatility reshapes the GDP growth distribution rather than simply shifting its center.

The remainder of the paper is organized as follows. Section 2 describes the data and estimation strategy. Section 3 presents the main results. Section 4 presents the counterfactual analysis. Section 5 explores potential transmission mechanisms. Section 6 concludes.

2.2 Empirical Framework

2.2.1 Data

I use quarterly U.S. GDP growth* from the Federal Reserve of St. Louis FRED database, covering 1975Q1 to 2024Q1 (197 observations). I obtain monthly data on the U.S. Crude Oil Composite Acquisition Cost by Refiners from the U.S. Energy Information Administration (EIA)†. Since oil prices are non-stationary, I use quarterly and monthly oil price growth rates.

I use the National Financial Conditions Index‡ (NFCI) from the Federal Reserve of St. Louis. The NFCI is a weekly§ index comprising the weighted average of 105 financial sector measures (money markets, debt and equity markets, and the banking system), each standardized relative to its sample mean and standard deviation.

*Downloaded from FRED: https://fred.stlouisfed.org/series/A191RL1Q225SBEA.
†Downloaded from EIA: https://www.eia.gov.
‡Downloaded from FRED: https://fred.stlouisfed.org/series/NFCI.
§The NFCI is converted to quarterly frequency by averaging weekly observations within each quarter.

Positive values indicate tighter-than-average financial conditions, implying a negative correlation with GDP growth.

The mechanism analysis incorporates four additional variables. I use GDP forecast dispersion from the Survey of Professional Forecasters (SPF)[¶] — specifically the dispersion measure $D2$ (75th minus 25th percentile of quarter-over-quarter GDP growth forecasts) — as a proxy for ambiguity (Ilut & Schneider, 2014). I include the University of Michigan Consumer Confidence Index[‖] as a measure of household sentiment. I also use the Capacity Utilization index[**] and the Industrial Production index (INDPRO)[††], a broad measure of output in manufacturing, mining, and utilities.

2.2.2 Quantile Estimation Approach

To estimate the future GDP growth conditional on oil price fluctuation, financial conditions, and current economic activity, I rely on the quantile estimator as proposed by Koenker and Bassett Jr (1978):

$$\hat{\beta}_\tau = argmin_{\beta_\tau \in \mathbb{R}^k} \{ \sum_{t=1}^{T} \tau \times \mathbf{1}_{y_{t+h} > x_t \beta_\tau} (y_{t+h} - x_t \beta_\tau) + (1-\tau) \times \mathbf{1}_{y_{t+h} < x_t \beta_\tau} (y_{t+h} - x_t \beta_\tau) \} \tag{2.1}$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function, y_{t+h} is the annualized average GDP growth rate between t and $t+h$, and x_t is the vector of conditioning variables (oil price fluctuations, NFCI, and current GDP growth), including a constant. The estimated value is the τ-th conditional quantile of y_{t+h} given x_t:

$$\hat{Q}_{y_{t+h}|x_t} (\tau | x_t) = x_t \hat{\beta}_\tau \tag{2.2}$$

For each period t, I estimate quantiles $\tau \in \{0.05, 0.10, \ldots, 0.95\}$, yielding 19 conditional quantile estimates per quarter, at horizons $h = 1$ and $h = 4$.

[¶] Available at: https://www.philadelphiafed.org/surveys-and-data/ngdp.
[‖] Downloaded from FRED: https://fred.stlouisfed.org/series/USACSCICP02STSAM.
[**] Downloaded from FRED: https://fred.stlouisfed.org/series/TCU.
[††] Downloaded from FRED: https://fred.stlouisfed.org/series/INDPRO.

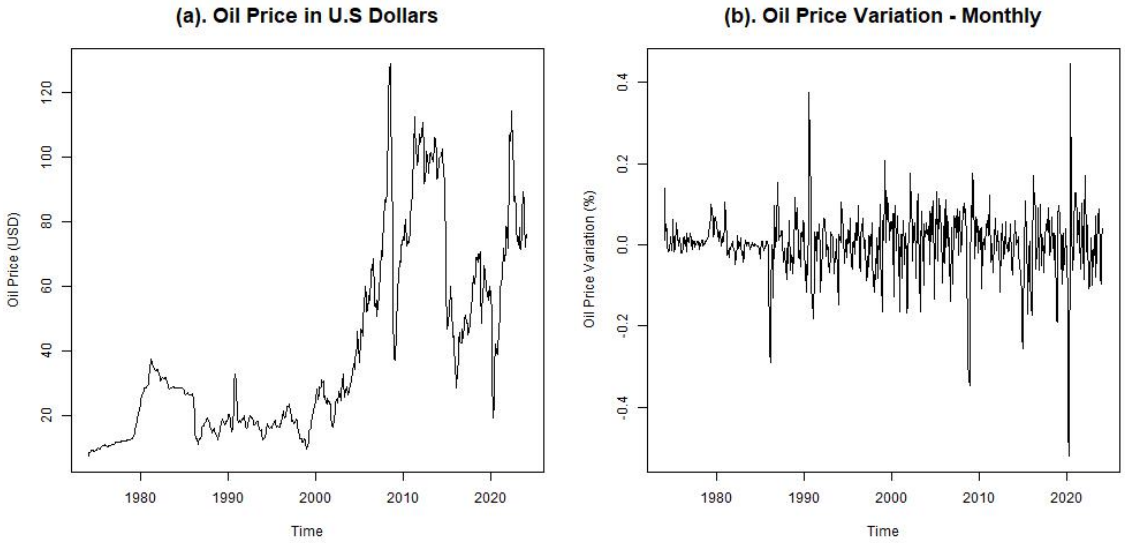


Figure 2.1: Oil Prices - From 1974 until 2024

Table 2.1: Data Summary

Variable	Proxy	Frequency	Source
Economic Output	U.S. GDP Growth	Quarterly	FRED database
Oil Prices	U.S. Crude Oil Composite	Monthly	U.S. EIA
Financial Conditions	NFCI	Quarterly	FRED database
Ambiguity	GDP Dispersion	Quarterly	SPF (Phil. Fed)
Consumer Confidence	Consumer Confidence Index	Quarterly	FRED database
Capacity Utilization	Capacity Utilization Index	Quarterly	FRED database
Industrial Production	INDPRO	Quarterly	FRED database

2.2.3 Quantile Estimation with Mixed Frequency Data

Oil price data are observed monthly, yielding multiple observations per GDP quarter. Not all lags, however, carry predictive information about future GDP growth.

To select the optimal lag structure, I follow Lima et al. (2020). The procedure has two steps:

1. *Soft-Thresholding (Lag Selection):* For each period t, I apply elastic net (EN) penalization (Friedman et al., 2010) to the mean GDP model to select the most predictive lag structure for monthly oil price fluctuations.

2. *Hard-Thresholding:* Given the selected lag structure, I retain only regressors with p-values ≤ 0.05 in the quantile regression.

Alternatively, the penalization in step 1 can be applied separately for each quantile τ and period t, allowing a distinct lag structure per quantile. Lima et al. (2020) show this flexibility improves out-of-sample performance at the cost of greater computational effort.

In summary, Model 1 uses quarterly oil price growth (end-of-quarter percent change). Model 2 uses monthly oil price lags selected via elastic net. Models 3 and 4 follow Adrian et al. (2019) (NFCI and current GDP) and an AR(1) benchmark, respectively. Table 2.2 summarizes all specifications.

2.2.4 Out-of-sample Evaluation

I evaluate the four models via a recursive out-of-sample forecasting exercise. Using 1975Q1–1994Q4 as the initial training sample (80 observations), I generate forecasts for 1995Q1 ($h = 1$) and 1995Q4 ($h = 4$). The training window then expands by one quarter at each step, with forecasts computed through 2024Q1.

At each iteration, I compute $\hat{y}_{t+h}$ across quantiles $\tau \in \{0.05, 0.10, \ldots, 0.95\}$:

$$\hat{y}_{t+h} = \hat{Q}_{y_{t+h}|x_t} = x_t \hat{\beta}_\tau \tag{2.3}$$

To assess out-of-sample performance, I follow Koenker and Machado (1999) and compute the out-of-sample Pseudo R^2, which compares the conditional quantile forecasts $\widehat{y}^i_{t+h,\tau}$, $i \in \{1,2,3\}$, to the AR(1) benchmark $\overline{y}_{t+h,\tau}$, as summarized in Table 2.2[‡‡].

The Pseudo $R^2_{OOS}(\tau)$ is defined as:

$$R^2_{OOS}(\tau) = 1 - \frac{\sum_{t=T_{in}}^{T-1} \rho_\tau(y^{obs}_{t+h} - \widehat{y}^i_{t+h,\tau})}{\sum_{t=T_{in}}^{T-1} \rho_\tau(y^{obs}_{t+h} - \overline{y}_{t+h,\tau})} \tag{2.4}$$

where T is the total number of periods, T_{in} is the number of in-sample periods, y^{obs}_{t+h} is the observed output in period $t+h$, $\widehat{y}^i_{t+h,\tau}$ is the τ-th quantile forecast from model $i = 1,2,3$, and $\overline{y}_{t+h,\tau}$ is the corresponding benchmark quantile forecast. The quantile loss function $\rho_\tau(\cdot)$ is defined as in Koenker and Bassett Jr (1978):

$$\mathcal{L}(u_t) = \rho_\tau(u_t) = u_t(\tau - I(u_t < 0)) \tag{2.5}$$

where u_t is the residual at period t. A positive $R^2_{OOS}(\tau)$ indicates that the alternative model has lower quantile loss than the benchmark at quantile τ.

For Model 2, the lag selection procedure of Lima et al. (2020) is applied within each recursive window. At each iteration t, the elastic net is estimated using only in-sample observations to determine the optimal lag structure for monthly oil price fluctuations; $\hat{y}_{t+h}$ is then computed using that structure. The training window expands by one quarter and the procedure repeats, ensuring that no out-of-sample information enters the lag selection at any point.

[‡‡]Model 4 in Table 2.2 (GDP-AR1) serves as the benchmark.

The summation goes from T_{in} until $T-1$, so the last period T can be evaluated with $T-1$'s inputs.

2.2.5 Conditional Future GDP Growth

The quantile regression coefficients from equation 2.2 yield a conditional quantile function. Following Azzalini and Capitanio (2003), I fit the estimated conditional quantiles to a skewed t-distribution to recover a probability density function:

$$f\left(y;\mu,\sigma,v,\alpha\right)=\frac{2}{\sigma}t\left(\frac{y-\mu}{\sigma};v\right)T\left(\alpha\frac{y-\mu}{\sigma}\sqrt{\frac{v+1}{v+\left(\frac{y-\mu}{\sigma}\right)^{2}}};v+1\right) \tag{2.6}$$

where $t(\cdot)$ and $T(\cdot)$ are the PDF and CDF of the Student-t distribution, μ is the location, σ the scale, v the degrees of freedom, and α the shape parameter.

For each period, the four parameters $\{\mu_t, \sigma_t, v_t, \alpha_t\}$ are chosen to minimize the squared distance between the estimated conditional quantile function $\hat{Q}_{y_{t+h}|x_t}(\tau|x_t)$ and the quantile function of the skewed t-distribution $F^{-1}(\tau;\mu_t,\sigma_t,v_t,\alpha_t)$ at $\tau = (0.05, 0.25, 0.50, 0.75, 0.95)$, from which I construct the conditional GDP growth density:

$$\{\hat{\mu}_{t+h},\hat{\sigma}_{t+h},\hat{v}_{t+h},\hat{\alpha}_{t+h}\}=\underset{\mu,\sigma,v,\alpha}{\arg\min}\sum_{\tau}\left(\hat{Q}_{y_{t+h}|x_t}\left(\tau|x_t\right)-F^{-1}\left(\tau;\mu_t,\sigma_t,v_t,\alpha_t\right)\right)^{2} \tag{2.7}$$

To evaluate the fit of the conditional GDP growth distribution, I apply the Berkowitz (Berkowitz, 2001) test, which assesses whether the predicted distribution matches the empirical data distribution.

2.2.6 Oil and GDP Downside Risk

The applied macroeconomics literature extends value-at-risk (VaR) to GDP growth through the growth-at-risk (GaR) concept, defined as the 5th percentile of the conditional GDP growth distribution:

$$GaR=GDP_{Growth}\in\tau\leq 0.05=\hat{Q}_{y_{t+h}|x_t}\left(0.05|x_t\right)=x_t\hat{\beta}_{(\tau=0.05)} \tag{2.8}$$

Complementary to GaR, I follow Adrian et al. (2019) and compute downside entropy, which compares the probability assigned to below zero GDP growth under the conditional density to that under the unconditional density.

Denoting $\hat{g}_{y_{t+h}}$ the unconditional density and $\hat{f}_{y_{t+h}|x_t}(y|x_t)$ the estimated skewed t-distribution, where $\hat{f}_{y_{t+h}|x_t}(y|x_t) = f(y; \hat{\mu}_{t+h}, \hat{\sigma}_{t+h}, \hat{v}_{t+h}, \hat{\alpha}_{t+h})$. The downside entropy L_t^D is:

$$L_t^D\left(\hat{f}_{y_{t+h}|x_t}; \hat{g}_{y_{t+h}}\right) = -\int_{-\infty}^{F^{-1}_{y_{t+h}|x_t}(0.5|x_t)} \left(\log \hat{g}_{y_{t+h}} - \log \hat{f}_{y_{t+h}|x_t}(y|x_t)\right) \hat{f}_{y_{t+h}|x_t}(y|x_t)\, \mathrm{d}y \tag{2.9}$$

where $F^{-1}_{y_{t+h}|x_t}(0.5|x_t)$ is the conditional median.

These two measures characterize the left-tail effects of oil price fluctuations on the conditional GDP growth distribution, assessing whether oil price fluctuations shift probability mass toward the left tail and raise the likelihood of non-positive growth outcomes. I also compare distributional moments — variance and kurtosis — across models to assess the impact of oil prices on macroeconomic uncertainty.

2.3 Results and Discussion

2.3.1 In-sample Quantile Estimates

I first estimate the unconditional GDP growth distribution over the full sample (1975Q1–2024Q1). Figure 2.2 plots the distribution using an Epanechnikov kernel (Epanechnikov, 1969).

In the conditional model, oil price fluctuations yield statistically significant coefficients only at the lowest quantiles ($\tau < 0.10$), with a negative sign indicating that extreme oil price movements amplify downside GDP risk. This effect is invisible in mean or median frameworks and only emerges at the distributional tails. Table 2.4 reports the quantile regression coefficients for the one-quarter-ahead horizon and

Table 2.2: Specification Summary

Model	Specification
1	Oil price fluctuation, financial conditions (NFCI), and current GDP (all variables with quarterly frequency).
2	Monthly oil price fluctuation filtered by elastic net, financial conditions (NFCI), and current GDP (NFCI and GDP still with quarterly frequency).
3	Financial conditions (NFCI) and current GDP (both with quarterly frequency).
4	Current GDP, which is equivalent to an AR(1) process and will serve as benchmark when estimating out-of-sample performance.

Table 2.3: Unconditional GDP Growth

τ	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
GDP Growth	−2.90	−0.70	0.50	1.10	1.50	2	2.20	2.30	2.50	3
τ	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.900	0.950	
GDP Growth	3.20	3.50	3.80	4	4.40	4.90	5.40	6.70	7.70	

*Unconditional GDP Growth is obtained by running quantiles regressions with intercept only.

Table 2.5 for the one-year-ahead horizon. The oil price variable shown is the quarterly growth rate without the elastic net filter.

When the elastic net procedure of Section 2.2.3 is applied over the full sample, it selects only the most recent month of oil price fluctuation, which is significant at the 5% level only for $\tau \leq 0.10$.

At the one-year-ahead horizon (Table 2.5), oil price coefficients are insignificant across all quantiles, consistent with a short-run effect of oil price fluctuations on GDP growth that dissipates over longer horizons.

2.3.2 Out of Sample Performance

Tables 2.6 and 2.7 report the pseudo-$R^2_{OOS}(\tau)$ for each model at the one-quarter-ahead and one-year-ahead horizons, respectively.

At the one-quarter-ahead horizon, oil price fluctuations improve forecast accuracy at the lower quantiles, particularly for $\tau \leq 0.10$, indicating that the oil models more accurately capture the Growth-at-Risk of future GDP growth.

When comparing the two oil models, the second model—which uses elastic net to select among monthly oil price lags (HF-EN) alongside NFCI and current GDP—outperforms the first model across almost all quantiles at the one-quarter-ahead horizon. Notably, Model 2 produces higher pseudo-R^2_{OOS} values not only at the lowest quantiles (e.g., $\tau = 0.05$: 12.77% vs. 11.22%; $\tau = 0.10$: 11.88% vs. 9.12%) but also across the middle of the distribution ($\tau = 0.40$–0.65), where Model 1 yields values close to zero while Model 2 consistently delivers gains of around 4%. Both oil models substantially outperform the financial conditions and current GDP model (NFCI/GDP, Model 3) for the left tail of the GDP distribution. These results motivate the density forecast analysis in the following section.

The pseudo-$R^2_{OOS}(\tau)$ in Table 2.7 confirm the results of Section 2.3.1: including oil price fluctuations yields no forecast accuracy gains at the one-year-ahead horizon. Most entries are negative across all three models, indicating that even the

Table 2.4: Quantile Regression Results - One Quarter Ahead

	GDP Growth			
	Lower Quantiles			
	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.15$	$\tau = 0.20$
Lag GDP	0.514^{***}	0.232^{***}	0.146^{**}	0.119^{*}
	(0.111)	(0.087)	(0.064)	(0.067)
NFCI	-1.369^{***}	-1.631^{***}	-1.667^{***}	-1.745^{***}
	(0.356)	(0.306)	(0.300)	(0.328)
Oil	-0.117^{***}	-0.040	-0.018	-0.013
	(0.032)	(0.024)	(0.014)	(0.013)
Intercept	-3.715^{***}	-1.025^{**}	-0.049	0.390
	(0.754)	(0.461)	(0.332)	(0.338)
	Median and Upper Quantiles			
	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$
Lag GDP	0.121^{*}	0.067	0.045	0.022
	(0.071)	(0.080)	(0.071)	(0.107)
NFCI	-1.717^{***}	-0.873^{**}	-0.930^{**}	0.297
	(0.342)	(0.416)	(0.398)	(0.772)
Oil	-0.017	-0.011	-0.015	-0.046
	(0.014)	(0.018)	(0.021)	(0.038)
Intercept	0.740^{**}	2.555^{***}	3.945^{***}	7.718^{***}
	(0.353)	(0.352)	(0.329)	(0.669)
Observations	197	197	197	197
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2.5: Regression Quantile Results - One Year Ahead

	GDP Growth			
	Lower Quantiles			
	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.15$	$\tau = 0.20$
Lag GDP	0.172	0.001	−0.083	−0.004
	(0.108)	(0.098)	(0.147)	(0.061)
NFCI	−1.854**	−1.632**	−1.287**	−1.115**
	(0.792)	(0.670)	(0.514)	(0.437)
Oil	0.055	−0.016	−0.011	−0.026
	(0.040)	(0.032)	(0.034)	(0.026)
Intercept	−3.174***	−0.347	0.472	0.911***
	(0.959)	(0.607)	(0.467)	(0.349)
	Median and Upper Quantiles			
	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$
Lag GDP	−0.043	−0.006	−0.013	0.001
	(0.086)	(0.068)	(0.047)	(0.050)
NFCI	−0.974**	−0.409	0.747	0.815**
	(0.401)	(0.566)	(0.475)	(0.328)
Oil	−0.016	−0.011	0.013	0.0005
	(0.026)	(0.023)	(0.015)	(0.014)
Intercept	1.492***	2.840***	4.658***	7.384***
	(0.363)	(0.370)	(0.330)	(0.354)
Observations	193	193	193	193
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2.6: Pseudo R2 OOS Results - One Quarter Ahead

τ	Oil/NFCI/GDP (1)	Oil(HF-EN)/NFCI/GDP (2)	NFCI/GDP (3)
0.05	11.220	12.770	−1.500
0.10	9.120	11.880	5.760
0.15	6.230	10.490	7.170
0.20	4.910	6.830	3.790
0.25	6.460	7.950	4.160
0.30	4.680	6.440	3.850
0.35	2.570	4.340	2.890
0.40	1.740	4.800	1.730
0.45	1.310	4.130	2.260
0.50	1.120	4.020	1.160
0.55	0.650	4.010	0.380
0.60	0.290	3.860	0.630
0.65	1.020	4.930	1.520
0.70	1.400	2.200	1.890
0.75	−2.130	0.050	−0.950
0.80	−0.400	0.690	−0.070
0.85	−0.740	−0.130	0.590
0.90	−1.790	−0.590	0.810
0.95	−40.290	−41.450	−1.260

Notes: Each entry reports the pseudo-$R^2_{OOS}(\tau)$ (in %) following Koenker and Machado (1999), computed over a recursive expanding window from 1995Q1 to 2024Q1 at the one-quarter-ahead horizon. A positive value indicates that the column model outperforms the GDP-AR1 benchmark (Model 4) at that quantile. **(1) Oil/NFCI/GDP**: quarterly oil price fluctuation, NFCI, and current GDP. **(2) Oil(HF-EN)/NFCI/GDP**: monthly oil price fluctuation filtered by elastic net, NFCI, and current GDP (NFCI and GDP at quarterly frequency). **(3) NFCI/GDP**: financial conditions (NFCI) and current GDP at quarterly frequency. Benchmark: current GDP (GDP-AR1).

Table 2.7: Pseudo R2 Out of Sample Results - One Year Ahead

τ	Oil/NFCI/GDP (1)	Oil(HF-EN)/NFCI/GDP (2)	NFCI/GDP (3)
0.05	−4.46	−3.96	−2.51
0.10	−2.97	−1.87	−0.04
0.15	−0.57	0.26	0.85
0.20	1.59	0.73	1.65
0.25	0.99	1.30	1.43
0.30	−0.04	0.21	0.73
0.35	−1	−1.28	−0.31
0.40	−2.22	−1.13	−0.53
0.45	−2.21	−2.16	−1.39
0.50	−1.12	−0.60	0.06
0.55	−0.97	0.23	0.82
0.60	−1.76	−1.11	−0.79
0.65	0.12	0.29	0.36
0.70	−1.26	−1.03	−0.74
0.75	−2.93	−2.92	−2.48
0.80	−1.22	−0.07	−0.02
0.85	−1.40	−0.35	−0.38
0.90	5.07	2.43	2.93
0.95	2.44	4.40	4.39

Notes: Each entry reports the pseudo-$R^2_{OOS}(\tau)$ (in %) following Koenker and Machado (1999), computed over a recursive expanding window from 1995Q4 to 2024Q1 at the one-year-ahead horizon. A positive value indicates that the column model outperforms the GDP-AR1 benchmark (Model 4) at that quantile.

parsimonious NFCI/GDP specification (Model 3) struggles to beat the GDP-AR1 benchmark at longer horizons. Accordingly, the density forecast analysis is restricted to the one-quarter-ahead horizon.

Lastly, I also estimated a fully quantile-specific elastic net, allowing a distinct lag structure for each τ and period t. This extension does not improve upon Models 1 or 2.

2.3.3 Density Forecast

Using the skewed t-distribution framework of Section 2.2.5, I construct one-quarter-ahead density forecasts for each period t using only information available up to t.

I compare three density forecasts: (i) GDP growth conditional on oil price fluctuations, financial conditions, and current GDP (Model 1); (ii) GDP growth conditional on financial conditions and current GDP (Model 3); and (iii) the AR(1) benchmark (Model 4).

Figure 2.3 shows the density forecast for 2008Q4, at the onset of the Great Recession. The oil model places substantially more mass to the left tail relative to the financial-conditions and AR(1) models, indicating that oil price information predicts a higher probability of negative growth at that point in time.

For 2024Q1 (Figure 2.4), the three density forecasts are more closely aligned, though the oil model retains a marginally heavier left tail. The time-varying nature of this effect is explored further in Section 2.4. Density forecasts for all periods are available in the online appendix.

Table 2.8 reports Berkowitz test p-values for each model, ranging from 0.04 to 0.06. These borderline p-values may reflect limited statistical power with 117 out-of-sample observations, as the Berkowitz test requires substantial data to detect misspecification in the variance and autocorrelation components (Berkowitz, 2001).

Both oil models yield similar density forecast curves; Model 1 is shown for brevity.

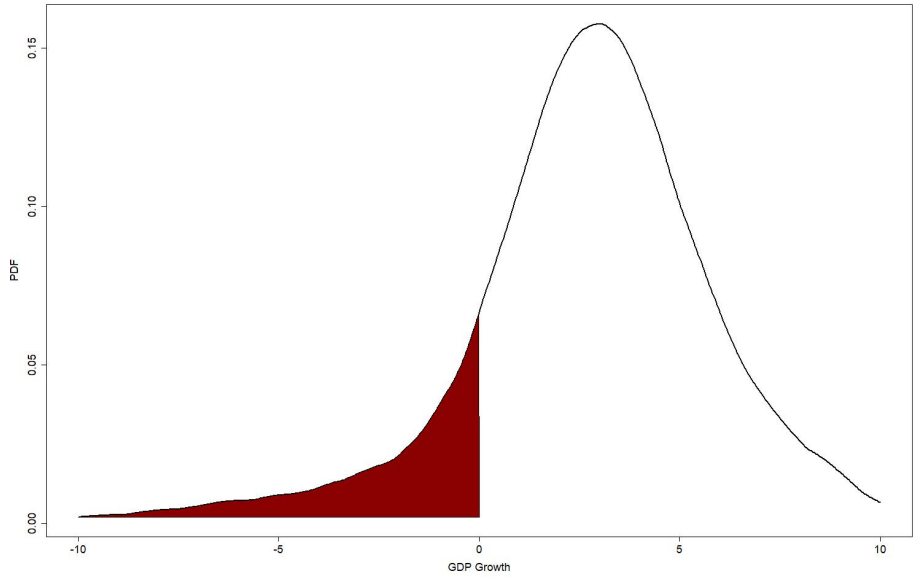


Figure 2.2: Unconditional GDP Growth Distribution

This figure shows the unconditional GDP growth distribution, estimated by (i) fitting a quantile regression with an intercept only and (ii) applying an Epanechnikov kernel to recover the density. The shaded area indicates negative growth.

Table 2.8: Berkowitz test statistics

	Models	P-Values (Berkowitz Test)
1	Oil(Quarterly)/NFCI/GDP	0.05
2	Oil(EN)/NFCI/GDP	0.04
3	NFCI/GDP	0.05
4	GDP(AR1)	0.06

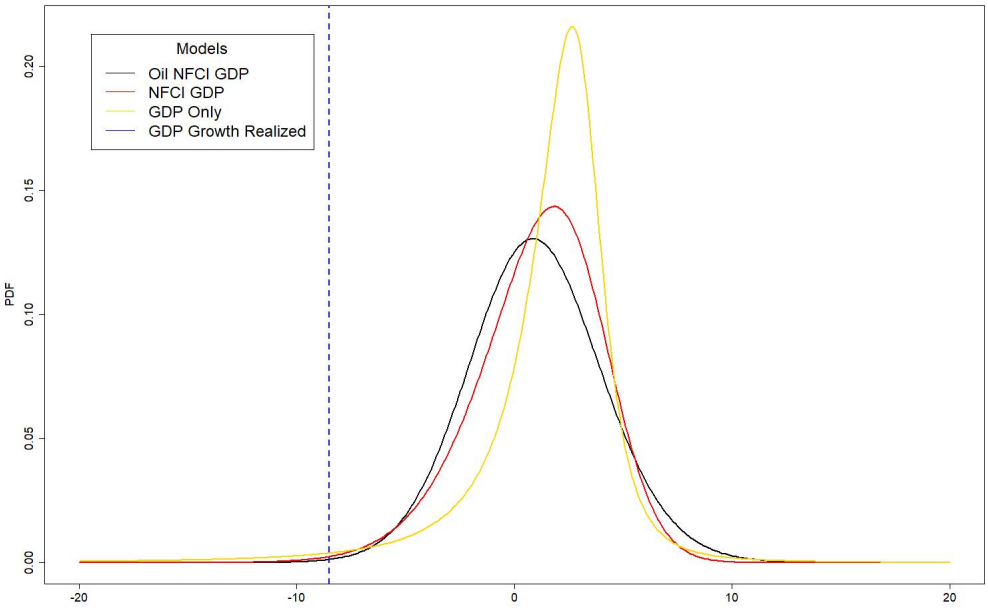


Figure 2.3: Conditional GDP Growth Distribution, One Quarter Ahead: 2008Q4

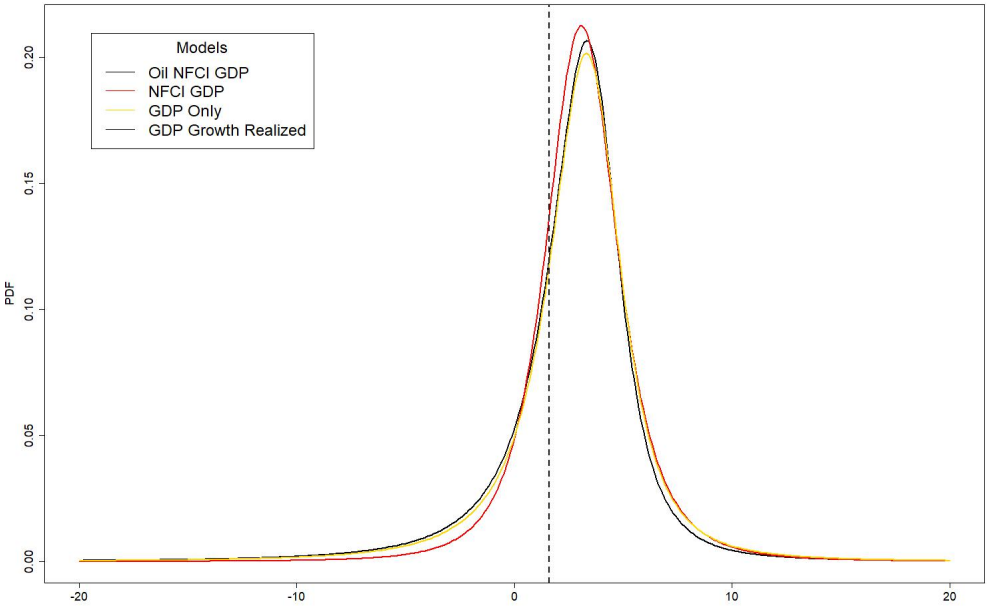


Figure 2.4: Conditional GDP Growth Distribution, One Quarter Ahead: 2024Q1

2.3.4 Vulnerability Measures

I report results for three measures of macroeconomic tail risk: GaR, left-side entropy, and density forecast variance.

GaR is the 5th-percentile conditional quantile forecast for each period t. Figure 2.5 plots its evolution over the sample. The oil model (Oil/NFCI/GDP) produces a more negative GaR on average than the financial-conditions model (NFCI/GDP), though the difference is not statistically significant. As established in Section 2.3.2, the oil models yield greater forecast accuracy at the lowest quantile, implying a more precise GaR estimate.

Left-side entropy shows no significant difference across models in the out-of-sample analysis. The gap widens in the in-sample analysis, as reported in Tables 2.9 and 2.10.

The GaR discrepancy between models has widened in recent years (Figure 2.5), and the oil price coefficient at $\tau = 0.05$ has strengthened over time (Figure 2.6), suggesting that the tail dependency of GDP growth on oil price fluctuations is a recent phenomenon. This may reflect structural changes in the U.S. energy sector, including the transition to net oil exporter status in 2019. Section 2.4 investigates this further via a counterfactual exercise.

In addition, I use density forecast variance as a proxy for macroeconomic uncertainty. The oil model (Oil/NFCI/GDP) produces 5% higher variance on average than the NFCI/GDP model in the out-of-sample analysis; this gap doubles in the in-sample analysis. While prior literature documents a declining macroeconomic sensitivity to oil price shocks (Hamilton, 1996; Blanchard and Gali, 2007; Kilian and Vigfusson, 2017), these results indicate that oil price fluctuations retain informational content about short-run GDP growth probabilities, particularly in the tails.

In-sample analysis refers to using the entire sample to estimate the density forecast, as opposed to only information available up to period t.

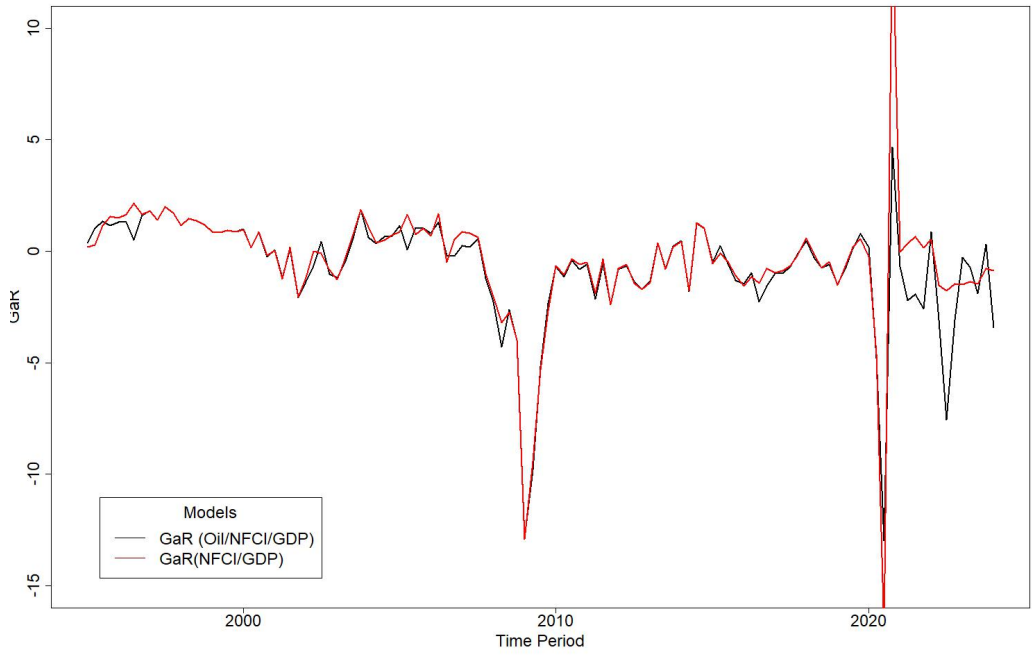


Figure 2.5: Growth-at-Risk measure - One quarter ahead

Table 2.9: Average of Vulnerability Measures - Out-of-Sample Analysis

Models	Growth-at-risk (mean)	Left-side Entropy	Variance
Oil/NFCI/GDP	-0.70	0.04	6.48
NFCI/GDP	-0.42	0.04	6.17

Table 2.10: Average of Vulnerability Measures - In-Sample Analysis

Models	Growth-at-risk (mean)	Left-side Entropy	Variance
Oil/NFCI/GDP	-2.50	0.15	10.65
NFCI/GDP	-1.67	0.12	9.02

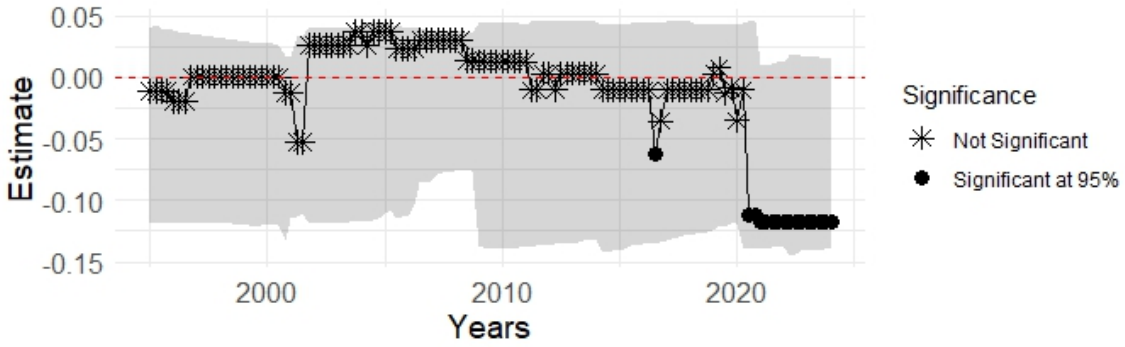


Figure 2.6: $\beta_{Oil|\tau=0.05}$ - Oil Estimate and Significance over time

2.3.5 The COVID-19 Pandemic and Structural Break

Figure 2.6 and Figure 2.5 suggest that the tail sensitivity of GDP growth to oil prices has intensified in recent years, raising the question of whether this reflects a discrete structural shift or whether the results are driven by a singular event. The COVID-19 pandemic is a natural candidate: the 2020 demand collapse sent oil prices briefly negative, and the subsequent recovery produced sharp price swings that may have altered the oil–growth relationship. I examine whether the significant oil price coefficients at the lower quantiles are driven entirely by the COVID-19 episode, or whether the tail dependence holds more broadly across the sample. I also test for structural change using the quantile structural break methodology of Qu (2008), which searches for breaks at each quantile separately and then jointly across all quantiles.

I first perform a leave-one-out exercise, excluding the four quarters of the most intense COVID-19 period (2020Q1–2020Q4) and re-estimating the quantile regression. Table 2.11 shows that the oil price coefficient at $\tau = 0.05$ remains negative and statistically significant, though at roughly half the magnitude of the full-sample estimate, while it loses significance at higher quantiles. This suggests that the tail dependence between oil prices and GDP growth is not driven entirely by the COVID-19 episode, but that the pandemic does amplify its magnitude. Figure 2.7 re-plots the rolling oil coefficient at $\tau = 0.05$ excluding the COVID period and confirms the same pattern, with the coefficient strengthening over time but at a more moderate level.

Having established that the tail dependence is not entirely driven by the COVID-19 episode, I now turn to the formal structural break test. Applying the quantile structural break methodology of Qu (2008), which searches for breaks at each quantile separately and then jointly across all quantiles, the test identifies a structural break

I perform the same exercise excluding the Global Financial Crisis period (2007Q3–2009Q4) and obtain a similar result: the oil price coefficient at $\tau = 0.05$ remains negative and statistically significant, but at roughly half the magnitude of the full-sample estimate.

Table 2.11: Regression Quantile Results - One Quarter Ahead (Excl. COVID 2020Q1–2020Q4)

	GDP Growth			
	Lower Quantiles			
	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.15$	$\tau = 0.20$
Lag GDP	0.163	0.257**	0.234**	0.187**
	(0.140)	(0.102)	(0.100)	(0.094)
NFCI	−2.311***	−1.568***	−1.815***	−1.637***
	(0.520)	(0.296)	(0.361)	(0.323)
Oil	−0.059**	−0.039	−0.005	−0.008
	(0.024)	(0.026)	(0.016)	(0.014)
Intercept	−1.841***	−1.075**	−0.441	0.187
	(0.498)	(0.466)	(0.434)	(0.379)
Observations	193	193	193	193
Note:				*p<0.1; **p<0.05; ***p<0.01

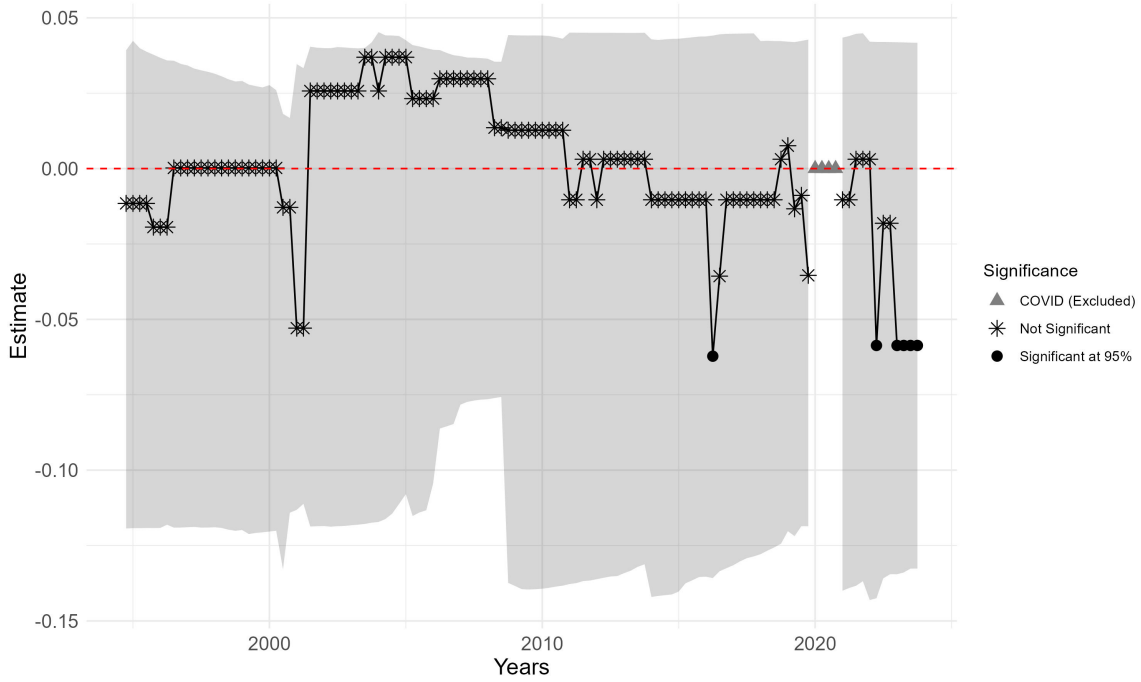


Figure 2.7: $\beta_{Oil|\tau=0.05}$ - Oil Estimate and Significance over time (Excl. COVID 2020Q1–2020Q4)

at 1989Q4. This result coincides with several candidate events: the 1986 oil price collapse reduced OPEC's influence, and the 1990 Gulf War triggered a sharp oil price spike. A limitation of this test is that it does not identify which regressor drives the break; the detected shift may reflect changes in financial conditions or technology unrelated to oil markets.

I re-estimate the quantile regressions separately for each sub-period. The oil price coefficient loses significance in both sub-samples, likely due to reduced statistical power with fewer observations. Since the break falls within the training sample (1975Q1–1995Q1), this does not invalidate the out-of-sample exercise, but it cautions against a time-invariant interpretation of the estimated coefficients.

More broadly, Figures 2.6 and 2.7 suggest that the dynamics governing the oil–growth relationship have shifted meaningfully in the post-2020 period, with the tail coefficient strengthening and stabilizing at a persistently negative level. Whether this reflects a permanent change in how oil price fluctuations transmit to downside growth risk — through energy transition pressures, supply chain reorganization, or altered monetary policy responses — remains an open question. Future work could apply a time-varying coefficient quantile regression framework to track these dynamics more formally, or extend the sample as post-pandemic data accumulates to assess whether the pattern is durable.

2.4 Counterfactual Analysis - CoVaR

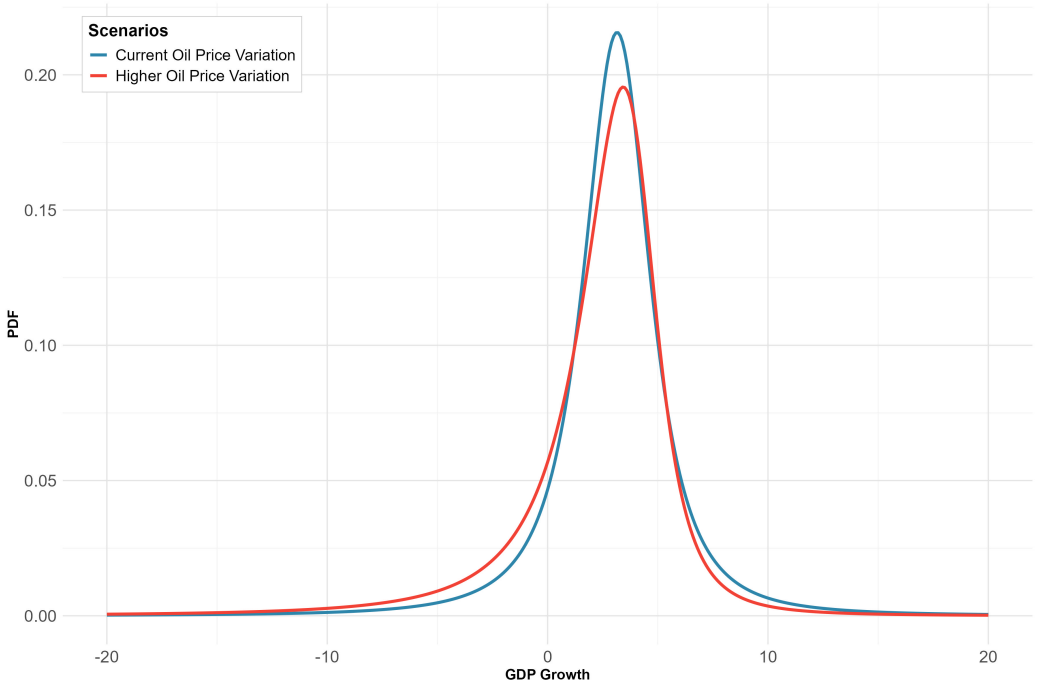
Following Adrian and Brunnermeier (2016), I extend the ΔCoVaR framework to a macroeconomic setting. ΔCoVaR is defined as the change in the value-at-risk of the financial system conditional on a particular institution being in distress relative to its median state; here, I apply the analogous logic to GDP growth, replacing the institution-level distress with an extreme oil price fluctuation. I estimate how a shift in oil price volatility from its median to the 97.5th percentile ($OPV_{\tau>0.975}$), holding all other variables fixed, affects the conditional GDP growth distribution.

Loria et al. (2023) estimate the impulse response of industrial production to oil price shocks, focusing on the 10th quantile. I extend this by estimating changes across the full conditional GDP growth density, including distributional moments such as variance and interquartile range.

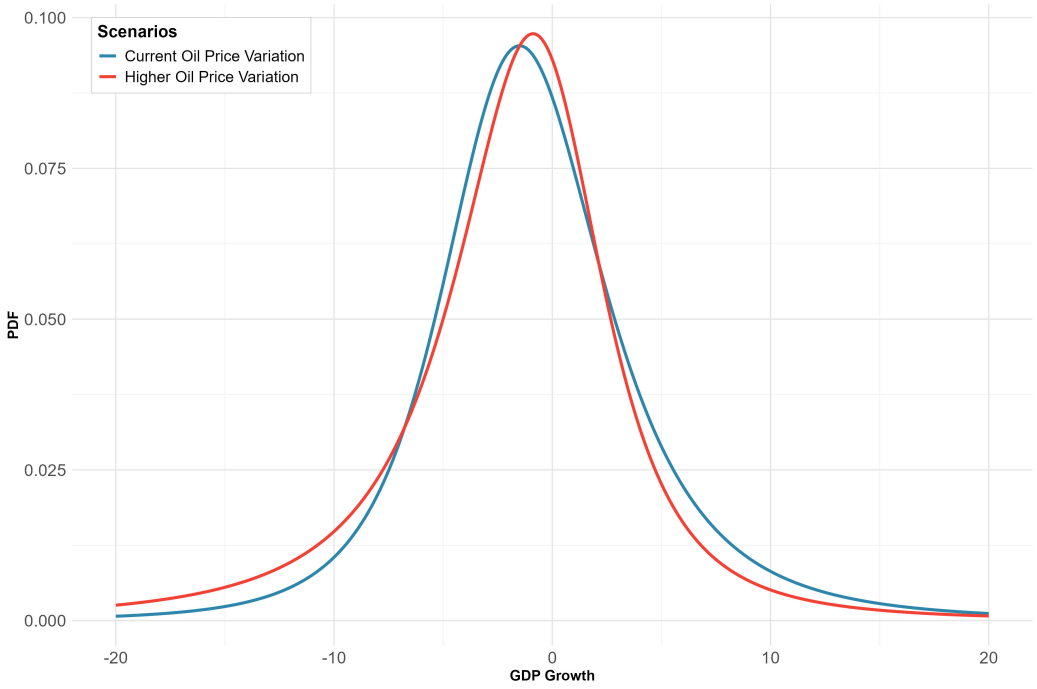
To assess how the effect of oil price variation varies across economic regimes, I consider three conditioning periods. The baseline is 2017Q1, when all conditioning variables were near their sample medians ($OPV_t = -1.22$, $NFCI = -0.4$, $y_t = 2.2$), representing normal economic and financial conditions. I also consider 2009Q1 as a recession scenario ($NFCI = 2.4$), and 1997Q4 as a booming conditions scenario ($NFCI = -0.65$, $y_t = 5.1$). In each case, I replace the observed oil price variation with the 97.5th-percentile value, holding all other variables fixed, and re-estimate the one-quarter-ahead density forecast. Panels A, B, and C of Figure 2.8 compare the baseline and counterfactual densities for each regime.

Table 2.12 reports distributional moments for each scenario across three economic regimes. Under normal conditions (Panel A), variance increases by approximately 26% under the high oil price scenario, indicating greater uncertainty about future economic output. The effect is more muted during the recession period (Panel B), while booming conditions (Panel C) show the largest proportional increase in variance.

Table 2.13 reports the tail risk measures. Under normal conditions (Panel A), GaR deteriorates from -1.87 to -4.64 — a 2.5-fold increase in absolute value, consistent with Loria et al. (2023)'s finding that the left tail is three to six times more sensitive than the median. The recession probability rises from 9.95% to 15.96%, an increase of approximately 60%. During the recession period (Panel B), the already-elevated GaR worsens further from -8.95 to -14.31, while the recession probability increases from 59.59% to 62.90%. Booming conditions (Panel C) show the largest proportional deterioration in tail risk, with GaR shifting from near zero (0.15) to -2.82 and recession probability rising from 4.27% to 10.96%.

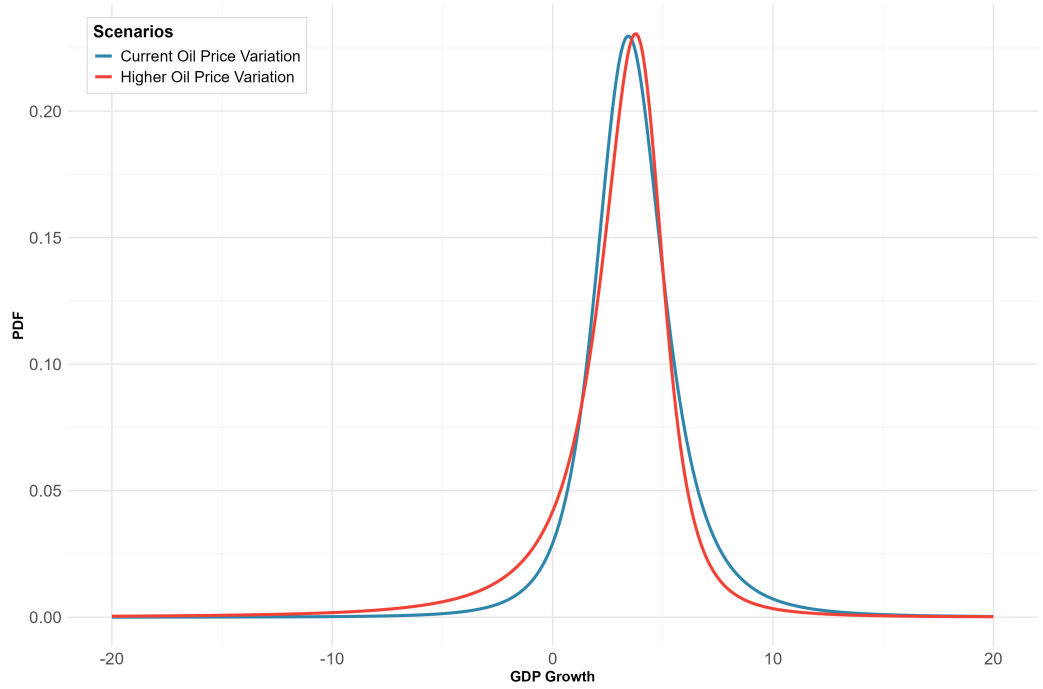


Panel A: Normal Economic and Financial Conditions (2017Q1)



Panel B: Recession Conditions (2009Q1)

Figure 2.8: Impact of Higher Oil Price Variation on Conditional GDP Growth



Panel C: Booming Conditions (1997Q4)

Figure 2.8: Impact of Higher Oil Price Variation on Conditional GDP Growth (continued)

Notes: Each panel compares the conditional density of one-quarter-ahead GDP growth under current oil price variation and a counterfactual scenario with oil price variation set to its 97.5th percentile, holding all other variables fixed. Panel A corresponds to normal economic and financial conditions (2017Q1; $NFCI = -0.4$; GDP growth = 2.2%). Panel B corresponds to recession conditions (2009Q1; $NFCI = 2.4$). Panel C corresponds to booming economic conditions (1997Q4; $NFCI = -0.65$; GDP growth = 5.1%). The shift in each distribution reflects changes in both tail risk and dispersion.

Table 2.12: CoVaR – Moments of the Density Forecast

Scenario	Mean	Variance	Interq. Range	Kurtosis
Panel A: Normal Economic and Financial Conditions (2017Q1)				
Median OPV	2.9893	10.0189	2.7092	10.4060
Higher OPV	2.3279	12.6420	2.7529	9.6654
Panel B: Recession Conditions (2009Q1)				
Median OPV	−0.7574	27.2053	5.8717	4.4926
Higher OPV	−1.5830	28.2920	5.9564	4.6719
Panel C: Booming Conditions (1997Q4)				
Median OPV	3.6953	6.2511	2.3021	9.0935
Higher OPV	2.8859	9.9511	2.3490	11.9078

Notes: This table reports the moments of the one-quarter-ahead conditional GDP growth density forecast under a counterfactual increase in oil price variation (OPV) to the 97.5th percentile of its historical distribution. "Median OPV" uses the actual oil price variation observed at the selected period; "Higher OPV" replaces it with the 97.5th percentile. Panel A corresponds to normal economic and financial conditions (2017Q1; NFCI = −0.4; GDP growth = 2.2%). Panel B corresponds to recession conditions (2009Q1; NFCI = 2.4). Panel C corresponds to booming economic conditions (1997Q4; NFCI = −0.65; GDP growth = 5.1%). IQR = Interquartile Range (Q75 − Q25).

Table 2.13: CoVaR – Tail Risk Measures

Scenario	Growth-at-Risk (GaR)	Left-Side Entropy	Recession Probability (%)
Panel A: Normal Economic and Financial Conditions (2017Q1)			
Median OPV	−1.8704	−0.0110	9.9478
Higher OPV	−4.6391	0.0239	15.9577
Panel B: Recession Conditions (2009Q1)			
Median OPV	−8.9457	0.9038	59.5880
Higher OPV	−14.3122	0.8935	62.8957
Panel C: Booming Conditions (1997Q4)			
Median OPV	0.1522	−0.0570	4.2669
Higher OPV	−2.8247	−0.0624	10.9568

Notes: This table reports tail risk measures from the one-quarter-ahead conditional GDP growth density forecast under a counterfactual increase in oil price variation (OPV) to the 97.5th percentile of its historical distribution. "Median OPV" uses the actual oil price variation observed at the selected period; "Higher OPV" replaces it with the 97.5th percentile. Growth-at-Risk (GaR) is the 5th percentile of the conditional forecast distribution. Left-Side Entropy compares the additional probability mass below zero GDP growth between the conditional model and the unconditional distribution. Recession Probability is the percentage of probability mass below zero GDP growth. Panel A corresponds to normal economic and financial conditions (2017Q1; NFCI = −0.4; GDP growth = 2.2%). Panel B corresponds to recession conditions (2009Q1; NFCI = 2.4). Panel C corresponds to booming economic conditions (1997Q4; NFCI = −0.65; GDP growth = 5.1%).

2.5 Evidence on the Mechanism

The preceding results show that oil price fluctuations raise near-term recession risk. This section examines the mechanisms underlying this relationship. I consider three channels that operate at different margins: an ambiguity channel, in which oil prices act as noisy signals about future fundamentals; an uncertainty channel, in which fluctuations depress consumer confidence and spending; and a real-options channel, in which uncertainty delays firms' investment and reduces capacity utilization. These channels are not mutually exclusive, but they generate distinct empirical predictions across the distribution of outcomes, allowing me to assess their relative importance.

A. Ambiguity channel. One interpretation of oil price fluctuations treats price movements as noisy signals about the future state of productivity—a process that is exogenous but stochastic. When agents observe large swings in oil prices, they must form beliefs about whether these movements reflect durable shifts in fundamentals or transitory noise. If agents perceive this signal as noisy, Ilut and Schneider (2014) show that ambiguity-averse agents—that is, agents who dislike situations where the probabilities themselves are uncertain—will place disproportionate weight on adverse states when oil price fluctuations are large. Crucially, this is not merely pessimism: when agents cannot pin down the odds, they act *as if* the worst case is more likely. This belief-updating then propagates through the economy via three reinforcing margins. Firms, facing an adversely skewed distribution of future returns, cut investment. Households, uncertain about future income, reduce durable consumption. Lenders, unwilling to extend credit against an uncertain collateral value, tighten financial conditions. Taken together, these responses raise the probability of recession endogenously—even if no "true" productivity decline ever materializes.

This channel predicts that oil price fluctuations should generate *disagreement* among professional forecasters, since agents with different priors or information sets will respond differently to ambiguous signals. To test this, I replace y_{t+h} with the interquartile range of GDP growth forecasts from the Survey of Professional

Forecasters (SPF) as the outcome variable, restricting attention to $h = 1$ (one quarter ahead). If the signal channel is operative, larger oil price fluctuations should be followed by elevated disagreement about future output growth.

Table 2.14 reports the estimated coefficient on oil price fluctuations across 9 quantiles of the forecast dispersion distribution. The results offer limited support for the ambiguity channel. Across most of the distribution—from the 5th through the 85th quantile—the estimated coefficients are statistically insignificant, indicating no systematic association between oil price fluctuations and forecast disagreement in normal times. The one exception is at the 90th quantile, where the coefficient is positive and statistically significant at the 5% level ($\hat{\beta}_{0.90} = 0.0181$, $p = 0.010$). This pattern is consistent with an asymmetric interpretation: oil price fluctuations do not routinely elevate disagreement, but they do shift the upper tail of the dispersion distribution. This is consistent with the baseline results: if oil price fluctuations are associated with a thicker left tail of future GDP growth, forecasters conditioning on different priors will disagree more sharply about downside outcomes, pushing up the right tail of the cross-sectional forecast distribution.

B. Uncertainty channel. The second mechanism treats oil price fluctuations as uncertainty shocks that erode consumer confidence. Large and unpredictable oil price movements are salient to households, who observe gasoline prices directly and frequently. When prices fluctuate sharply, households face genuine uncertainty about their future purchasing power and real income. Agents respond to an increase in uncertainty by adopting a wait-and-see posture: deferring durable purchases, building precautionary savings, and reducing discretionary spending. Because durable consumption is lumpy and partially irreversible, these confidence effects can be persistent and quantitatively significant even when the underlying oil price shock is transitory.

To test this channel, I use the growth rate of the University of Michigan Consumer Sentiment Index as the outcome variable, with the same conditioning variables as in the baseline specification ($h = 1$, one quarter ahead). If oil price fluctuations transmit

to the macroeconomy through a confidence channel, larger fluctuations should be associated with lower sentiment growth across the distribution—and in particular at the lower quantiles, where adverse outcomes are concentrated.

Table 2.15 reports the estimated coefficient on oil price fluctuations across 9 quantiles of the consumer confidence growth distribution. The results provide considerably stronger support for the uncertainty channel than for the ambiguity channel documented above. The oil price coefficient is negative and statistically significant at the 5% level or better across virtually the entire distribution—from the 5th quantile ($\hat{\beta}_{0.05} = -0.093$, $p < 0.01$) through the 95th quantile ($\hat{\beta}_{0.95} = -0.149$, $p < 0.05$).

Two additional features of the results are noteworthy. First, the magnitude of the coefficient is broadly stable across quantiles, ranging from approximately -0.09 to -0.16 in absolute value. This stability implies that oil price fluctuations shift the entire distribution of confidence growth downward rather than compressing or stretching it—a location shift rather than a scale effect. Second, the coefficients tend to be somewhat larger in magnitude at the upper quantiles (0.80–0.90), suggesting that oil price fluctuations are particularly effective at curtailing episodes of strong sentiment growth. In other words, fluctuations not only depress confidence in bad times but also cap it in good times, consistent with an asymmetric dampening of the confidence cycle. Taken together, these results suggest that the confidence channel is the more empirically relevant transmission mechanism: oil price fluctuations systematically undermine consumer sentiment, which in turn suppresses household spending and raises the probability of recession.

C. Real-options channel. As an extension of the uncertainty channel, Bloom (2009) proposes that uncertainty shocks affect the economy not only through households but also through firms' investment decisions. Because investment is partially irreversible, firms adjust capital only when the gap between desired and actual capital crosses a threshold, remaining inactive otherwise. An increase in uncertainty—proxied here by large oil price fluctuations—widens this inaction

Table 2.14: Oil Price Coefficient on GDP Forecast Dispersion Across Quantiles

τ	$\hat{\beta}_\tau$	S.E.	p-value
0.05	-0.0006	(0.0045)	0.902
0.10	-0.0017	(0.0055)	0.765
0.15	0.0009	(0.0058)	0.873
0.25	0.0011	(0.0062)	0.859
0.50	0.0047	(0.0056)	0.407
0.75	0.0065	(0.0057)	0.258
0.85	0.0092	(0.0077)	0.235
0.90	0.0181^{**}	(0.0070)	0.010
0.95	0.0211	(0.0141)	0.136

Note: Each row reports the estimated coefficient on oil price fluctuation (Δoil_t) from a quantile regression of SPF GDP forecast dispersion on oil price fluctuation, the National Financial Conditions Index (NFCI), and current GDP growth, estimated at quantile τ. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 2.15: Oil Price Coefficient on Consumer Confidence Growth Across Quantiles

τ	$\hat{\beta}_\tau$	S.E.	p-value
0.05	-0.0935^{***}	(0.0320)	0.004
0.10	-0.0816^{***}	(0.0293)	0.006
0.15	-0.1052^{***}	(0.0288)	0.000
0.25	-0.1074^{***}	(0.0341)	0.002
0.50	-0.1033^{**}	(0.0468)	0.028
0.75	-0.0948^{*}	(0.0523)	0.071
0.85	-0.1570^{***}	(0.0546)	0.004
0.90	-0.1659^{***}	(0.0591)	0.005
0.95	-0.1490^{**}	(0.0670)	0.027

Notes: Each row reports the estimated coefficient on oil price fluctuation (Δoil_t) from a quantile regression of University of Michigan Consumer Sentiment Index growth on oil price fluctuation, the National Financial Conditions Index (NFCI), and current GDP growth, estimated at quantile τ. Standard errors (in parentheses) are kernel-based following Koenker and Bassett Jr (1978). ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

region, inducing firms to postpone capital adjustment and reducing aggregate factor utilization.

I test this mechanism using total capacity utilization (TCU) from FRED as the outcome variable. I include industrial production growth as an additional control for two reasons: first, to account for the direct cyclical relationship between output and capacity utilization; and second, to partial out positive demand shocks in which higher oil prices reflect strong economic performance rather than uncertainty. If the real-options channel is operative, oil price fluctuations should be negatively associated with capacity utilization, particularly at the lower quantiles.

Table 2.16 reports the results. The oil price coefficient is negative throughout the distribution but statistically significant only at the lower tail: at the 5th quantile ($\hat{\beta}_{0.05} = -0.056$, $p = 0.026$), and marginally at the 10th ($p = 0.051$) and 15th ($p = 0.091$). From the 20th quantile onward, the coefficients are small and insignificant.

This left-tail concentration mirrors the baseline GDP results and is consistent with the real-options mechanism: oil price fluctuations widen firms' inaction regions and compress the lower tail of capacity utilization, without affecting normal or high-utilization episodes. The evidence is directionally consistent with the theory, but narrower in scope than the confidence channel, which operates across the full distribution.

Taken together, the evidence suggests that multiple mechanisms simultaneously transmit oil price fluctuations to macroeconomic risk. The ambiguity channel receives limited support, operating primarily in the upper tail of forecast disagreement. The real-options channel is present but concentrated in the lower tail of capacity utilization, indicating effects on firms' downside adjustment margins. In contrast, the uncertainty channel operating through consumer sentiment exhibits a broad and robust impact across the entire distribution. This pattern suggests that while firm-side responses contribute to the amplification of downside risk, the primary transmission mechanism likely operates through households. Oil price fluctuations

act as salient signals of uncertainty that systematically depress consumer confidence, thereby reducing demand and increasing the likelihood of recession.

2.6 Discussion

This paper reexamines the relationship between oil price fluctuations and economic activity through the lens of asymmetric GDP growth distributions. Using quantile regression methods, I show that oil price fluctuations are a significant predictor of downside macroeconomic risk at short horizons.

The core finding is that oil price fluctuations disproportionately affect the left tail of the GDP growth distribution. At the one-quarter-ahead horizon, oil price variation is negatively and significantly associated with GDP growth for $\tau < 0.10$, with no corresponding effects at the median or upper quantiles. These results are corroborated out-of-sample: models incorporating oil price fluctuations improve forecast performance for lower quantiles and improve GaR forecast accuracy by approximately 12.77%. At longer horizons, however, these effects dissipate, consistent with the view that oil price fluctuations primarily operate as short-run risk factors.

A distributional analysis reveals that periods of heightened oil price volatility are associated with thicker left tails in the conditional GDP growth distribution, indicating elevated downside risk. The counterfactual analysis reinforces this interpretation: large oil price fluctuations substantially amplify macroeconomic tail risk, increasing both the likelihood and severity of adverse outcomes, as well as overall uncertainty.

The evidence on mechanisms points to a common underlying force: uncertainty. While the ambiguity channel finds limited support and the real-options channel operates primarily in the lower tail of firm activity, the consumer confidence channel exhibits a strong and pervasive response across the entire distribution. This suggests that oil price fluctuations transmit to the macroeconomy primarily by depressing

household sentiment, with firm-side adjustments reinforcing the contractionary effects in adverse states.

Whether these findings generalize across other sources of macroeconomic shocks and evolve over time remains an open question. Characterizing the dynamic response of the full GDP growth distribution to structurally identified shocks is a promising direction for future research.

Table 2.16: Oil Price Coefficient on Total Capacity Utilization Across Quantiles

τ	$\hat{\beta}_\tau$	S.E.	p-value
0.05	-0.0560^{**}	(0.0249)	0.026
0.10	-0.0570^{*}	(0.0291)	0.051
0.15	-0.0564^{*}	(0.0332)	0.091
0.25	-0.0558	(0.0390)	0.154
0.50	-0.0075	(0.0361)	0.835
0.75	-0.0210	(0.0294)	0.477
0.85	-0.0273	(0.0265)	0.304
0.90	-0.0325	(0.0251)	0.197
0.95	-0.0115	(0.0246)	0.640

Notes: Each row reports the estimated coefficient on oil price fluctuation (Δoil_t) from a quantile regression of total capacity utilization on oil price fluctuation, the National Financial Conditions Index (NFCI), current GDP growth, and industrial production growth, estimated at quantile τ. The inclusion of industrial production growth as an additional control accounts for the direct cyclical relationship between output and capacity utilization. Standard errors (in parentheses) are kernel-based following Koenker and Bassett Jr (1978). ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

References

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, *109*(4), 1263–1289.

Adrian, T., & Brunnermeier, M. K. (2016). Covar. *The American Economic Review*, *106*(7), 1705.

Aikman, D., Bridges, J., Hacioglu Hoke, S., O'Neill, C., & Raja, A. (2019). Credit, capital and crises: A gdp-at-risk approach.

Azzalini, A., & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *65*(2), 367–389.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of business & economic statistics*, *19*(4), 465–474.

Blanchard, O. J., & Gali, J. (2007). The macroeconomic effects of oil shocks: Why are the 2000s so different from the 1970s?

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623–685.

Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, *14*(1), 153–158.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*, 1–22.

Giglio, S., Kelly, B., & Pruitt, S. (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics*, *119*(3), 457–471.

Hamilton, J. D. (1996). This is what happened to the oil price-macroeconomy relationship. *Journal of monetary economics*, *38*(2), 215–220.

Ilut, C. L., & Schneider, M. (2014). Ambiguous business cycles. *American Economic Review*, *104*(8), 2368–2399.

IMF. (2017). *Global financial stability report, october 2017: Is growth at risk?* International Monetary Fund.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, *95*(1), 161–182.

Kilian, L., & Vigfusson, R. J. (2017). The role of oil price shocks in causing us recessions. *Journal of Money, Credit and Banking*, *49*(8), 1747–1776.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 33–50.

Koenker, R., & Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, *94*(448), 1296–1310.

Lima, L. R., Meng, F., & Godeiro, L. (2020). Quantile forecasting with mixed-frequency data. *International Journal of Forecasting*, *36*(3), 1149–1162.

Loria, F., Matthes, C., & Zhang, D. (2023). Assessing macroeconomic tail risk. *Available at SSRN 4002665*.

Qu, Z. (2008). Testing for structural change in regression quantiles. *Journal of Econometrics*, *146*(1), 170–184.

Sui, J., Lv, W., Gao, X., & Koedijk, K. G. (2024). China's gdp-at-risk: Real-time monitoring, risk tracing, and macroeconomic policy effects. *Journal of International Money and Finance*, *147*, 103150.

Chapter 3

Connected and Distracted: The Impact of High-Speed Internet on ADHD Outcomes in the U.S.

Abstract

I analyze the relationship between high-speed internet connectivity and ADHD-related outcomes in the United States using county-level data from 2008 to 2023. Exploiting variation in the timing and intensity of broadband adoption, I estimate two-way fixed effects models of the effect of high-speed internet penetration on county-level ADHD medication demand. In the main specification, which uses the Federal Communications Commission (FCC) internet penetration scale (0–5), each increase in broadband adoption tier is associated with a 0.85 percent increase in total ADHD medication demand, with amphetamine-based medications showing the highest increase at 1.65 percent. These effects are more pronounced in smaller counties. I also find some evidence of increased special education enrollment, though these estimates are less consistent across specifications. Using the American Time Use Survey, I find suggestive evidence that youth, especially boys, spend more time on computers as high-speed connectivity expands, although these estimates are imprecise. Overall, the results suggest that the expansion of digital infrastructure may have contributed to rising ADHD-related treatment demand.

3.1 High-Speed Internet Evolution and Attention-Deficit Disorders

The proliferation of high-speed internet has fundamentally reshaped how we allocate our time and conduct daily activities, with internet-connected devices becoming increasingly present across schools, homes, and workplaces. This digital transformation has created opportunities and challenges for human development. While research demonstrates that internet access can improve educational outcomes and enhance human capital formation (Dettling et al., 2018; Grimes & Townsend, 2018; Zuo, 2021), emerging evidence also reveals concerning negative impacts on health and human development (Arenas-Arroyo et al., 2025; Colombo & Failache, 2023; Jain & Stemper, 2025). Early medical studies associate increased screen exposure and hyperstimulation with difficulties in concentration and learning disabilities such as Attention Deficit Hyperactivity Disorder (ADHD), with children exposed to more television and video games showing higher rates of attention problems (Swing et al., 2010). Additional research has documented negative effects, including increased sedentary behavior, disrupted sleep patterns, and reduced face-to-face social interactions (Braghieri et al., 2022; Geraci et al., 2022; Twenge et al., 2018). As individuals continue to integrate digital technologies into every aspect of daily life, understanding these dual effects and their underlying mechanisms becomes critical for informing evidence-based policies and interventions.

In this paper, I document how access to high-speed internet affects the demand for ADHD treatment drugs and special education enrollment. Specifically, I leverage the expansion of high-speed internet connections (equal to or greater than 10 Mbps) across U.S. households as a quasi-natural experiment to evaluate its impact on ADHD medication distribution and special education enrollment at the county level. The analysis draws on comprehensive data covering all U.S. counties from 2008 to 2023, allowing for a robust examination of how the growth in high-speed

internet connections affects both the prevalence of attention-related disorders and the corresponding allocation of educational resources.

Understanding the relationship between high-speed internet access and ADHD outcomes is relevant given the long-term consequences of attention disorders on life trajectories, including employment, earnings and social interactions. High-speed internet enabled immersive streaming and instant content consumption that dial-up made impossible, creating a more stimulating digital environment that may impair sustained attention. Fletcher (2014) demonstrates that childhood ADHD significantly impairs adult labor market outcomes, with affected individuals experiencing reduced employment rates, lower wages, and decreased occupational attainment throughout their careers. Similarly, Colombo and Failache (2023) provide evidence that exposure to high-speed internet during early childhood development can have lasting negative effects on cognitive and behavioral outcomes, suggesting early digital exposure can produce deficits that persist into adulthood. Arenas-Arroyo et al. (2025) find that high-speed internet availability significantly widened the gender gap in adolescent mental health outcomes in Spain, with particularly pronounced negative effects on female teenagers' hospital admissions for mental health disorders. These findings highlight the importance of recognizing that the increased stimulus and constant availability of engaging content enabled by high-speed internet, particularly during early developmental stages, can have detrimental effects on children's long-term well-being, mental health, and economic prospects.

I draw on several publicly available data sources to analyze the relationship between broadband connectivity and ADHD-related outcomes over the period from 2008 to 2023. First, I utilize broadband connection data from the Federal Communications Commission (FCC) to track high-speed internet connectivity across U.S. counties. To measure ADHD prevalence and treatment, I employ the Drug Enforcement Administration's Automation of Reports and Consolidated Ordering System (DEA-ARCOS) dataset to track substances commonly used to treat ADHD at the county level. Finally, I incorporate educational data from the Common Core of

Data (CCD) maintained by the National Center for Education Statistics for special education funding distribution and enrollment patterns.

To identify the effects of high-speed internet expansion on the demand for ADHD medication and special education outcomes, I estimate two-way fixed effects models to assess the impact of high-speed internet penetration, using variation in the timing of broadband adoption across geographic units. I complement this analysis using difference-in-differences imputation as proposed by Borusyak et al. (2024), which addresses the negative weighting problem arising from heterogeneous treatment effects in staggered adoption settings, while still relying on the parallel trends assumption to validly impute counterfactual outcomes. These estimation methods enable me to separate the effects of broadband connectivity from confounding factors such as local economic conditions or demographic characteristics that might simultaneously influence both internet adoption and the outcomes of interest.

I find that greater high-speed internet penetration is associated with higher demand for ADHD medications. In the main specification, which uses the Federal Communications Commission (FCC) internet penetration scale (0–5), each increase in broadband adoption tier is associated with a 0.85 percent increase in total ADHD medication demand, with amphetamine-based medications showing the largest increase at 1.65 percent. Across alternative specifications, the estimated effects remain positive and generally range from 0.35 percent to 1.84 percent, depending on the treatment definition and sample restriction. These effects are more pronounced in smaller counties. I also find some evidence of increased special education enrollment, though these results are less robust than those for ADHD medications. Additional robustness checks suggest that the main findings are not driven by differential pre-trends or obvious confounding factors.

To explore potential mechanisms, I use data from the American Time Use Survey to show that high-speed internet adoption from 2008 to 2023 coincided with substantial increases in computer time among young males aged 15-20 years, whose weekly computing and gaming rose from approximately 7.5 hours in 2008 to

17.5 hours in 2023. In contrast, adult men aged 31-55 years exhibited relatively stable computer usage throughout this period. High-speed connectivity enabled very different forms of technology use—such as immersive multiplayer gaming and HD video streaming—that were not feasible with basic internet access during the earlier 2000s when information about ADHD was already widely available.

These findings contribute to the existing literature on technology's impact on human development and mental health. The most closely related works by Arenas-Arroyo et al. (2025), Donati et al. (2025), and Golin (2022), all of whom exploit geographic variation in broadband rollout to identify negative effects on adolescent mental health. My paper differs in three important respects: (i). I focus specifically on ADHD and attention-related outcomes rather than general mental health; (ii). I use pharmaceutical distribution data from DEA-ARCOS rather than hospital records or surveys, which overcomes survey credibility concerns and captures cases not severe enough to require hospitalization; and (iii). I provide county-level evidence for the full United States over a fifteen-year panel. In addition, several working papers also speak directly to the mechanisms I document: Colombo and Failache (2023) find negative effects of high-speed internet on early childhood cognitive development, Churchill and Johnson (2026) examine broadband and adolescent mental health in the U.S., and Caetano et al. (2025) provide evidence that video watching specifically harms children's skills. My findings join this growing body of evidence that expanded digital access carries developmental costs, particularly during childhood and adolescence.

From a policy perspective, this work has implications for federal broadband infrastructure programs and their unintended consequences. The Connect America Fund (CAF) has allocated over $10 billion over seven years to expand broadband infrastructure to underserved rural areas through various programs (Connect America Fund Phase II, Rural Digital Opportunity Fund, Alternative Connect America Cost Model). My findings are consistent with concerns in the medical literature (Danielson et al., 2022, 2024): smaller and rural counties—precisely those that received the most substantial CAF investments—exhibited more pronounced increases

in ADHD-related healthcare demand, if expanded digital access disproportionately affected ADHD prevalence in these communities, this points to a need for more coordinated approaches to digital infrastructure development that account for potential developmental and health costs alongside connectivity benefits.

The remainder of this paper is organized as follows. Section 2 describes the data sources and outlines the empirical strategy and identification approach. Section 3 presents the main results, and Section 4 examines alternative specifications and explores heterogeneity in treatment effects. Section 5 discusses potential mechanisms driving the observed effects, and Section 6 concludes with a summary of findings and discussion.

3.2 Empirical Framework

3.2.1 Data

This study draws on several publicly available data sources to analyze the relationship between broadband connections and ADHD-related outcomes over the main analysis period from 2008 to 2023. To measure broadband connectivity, I use connection data from the Federal Communications Commission (FCC) to track high-speed internet connections across U.S. counties. Connections are measured at the county level by tiers from 0 to 5, where tier 1 represents fewer than 200 connections per 1,000 households and tier 5 represents more than 800 connections per 1,000 households. Importantly, these figures reflect connections in use rather than connections available per household. I use connections of 10 Megabits per second or greater to designate an area as having high-speed connectivity. This represents the minimum reasonable speed that allows users to engage in online streaming and play online games, and also helps resolve compatibility issues with FCC data across different reporting periods.

To capture ADHD-related outcomes, I employ the Drug Enforcement Administration's Automation of Reports and Consolidated Ordering System (DEA-ARCOS)

dataset, which registers all controlled substances dispensed since 2006 at the three-digit zip code level. Specifically, I track substances commonly used to treat ADHD: amphetamine (Adderall), methylphenidate (Ritalin), and lisdexamfetamine (Vyvanse). I use the Department of Housing and Urban Development (HUD) Crosswalk files to transpose these pharmaceutical dispensing values to the county level. The HUD Crosswalk provides, for each zip code, the share of its residential population residing in each corresponding county. Since three-digit zip codes do not map cleanly onto county boundaries, I apply these population weights to allocate each zip code's dispensing totals proportionally across the counties it overlaps. DEA-ARCOS records the total grams of each substance shipped to DEA-registered pharmacies and hospitals rather than to final users; it therefore reflects the volume of medications received by dispensing outlets rather than individual-level consumption. Accordingly, I use county-level ADHD medication dispensing volumes as a proxy for ADHD prevalence across counties and years.

From the Common Core of Data (CCD) maintained by the National Center for Education Statistics, I obtain information on special education enrollment (Individualized Education Program) across school districts, which I aggregate to the county level. There are 13 disability categories recognized under the Individuals with Disabilities Education Act (IDEA) for an Individualized Education Program (IEP): autism spectrum disorder, deaf-blindness, deafness, emotional disturbance, hearing impairment, intellectual disability, multiple disabilities, orthopedic impairment, other health impairment, specific learning disability, speech or language impairment, traumatic brain injury, and visual impairment.

ADHD is explicitly included as a condition that can make a child eligible for special education services under the "Other Health Impairment" (OHI) category of IDEA. However, not all children with ADHD qualify for special education services, as the condition must adversely affect educational performance. Children with ADHD may also qualify under other categories, including specific learning disability or emotional disturbance, depending on their unique characteristics and educational needs. Among

students receiving special education services, specific learning disability is by far the most common category, representing approximately 32% of all students with IEPs, while other health impairments account for 15%. This distribution suggests that students with ADHD may be identified under multiple disability categories depending on their primary educational needs and the specific manifestation of their symptoms in the academic environment (CHADD, 2025; National Center for Education Statistics, 2023).

Additionally, I use data from the American Community Survey (ACS) at the county level for socio-demographic characteristics (population, income, education).

3.2.2 Background and National Trends

A. High-Speed Internet Connections. Figure 3.1 illustrates the expansion of high-speed internet access across U.S. counties from 2008 to 2023. The data shows a substantial increase in internet connectivity, with the population tier (representing the percentage of the households with access to high-speed internet) rising from approximately 1.2 (less than 20% coverage) in 2008 to 4.5 (80% or higher coverage) by 2023. The geographic distribution of this expansion is shown in the county-level maps presented in Figure 3.2, which reveal how internet access improved from sparse coverage in rural areas in 2008 to near-universal coverage by 2023. This temporal and spatial variation supports the use of an estimator with time and area fixed effects.

B. ADHD National Trends. ADHD diagnoses among U.S. children have risen sharply over the past two decades. Data from the Centers for Disease Control and Prevention show that ADHD diagnoses among children aged 5-17 years rose from 6.9% in 1998-2000 to 11.4% in 2022 (Akinbami et al., 2011; Danielson et al., 2024), representing an approximate 65% increase. Rising ADHD diagnoses extend beyond childhood into young adulthood. Among adults aged 18-44 years, National Health Interview Survey data reveal that diagnosed ADHD prevalence increased from 0.9% in 1997-1998 to 4.0% in 2017-2018 (Chung et al., 2019). This fourfold increase was

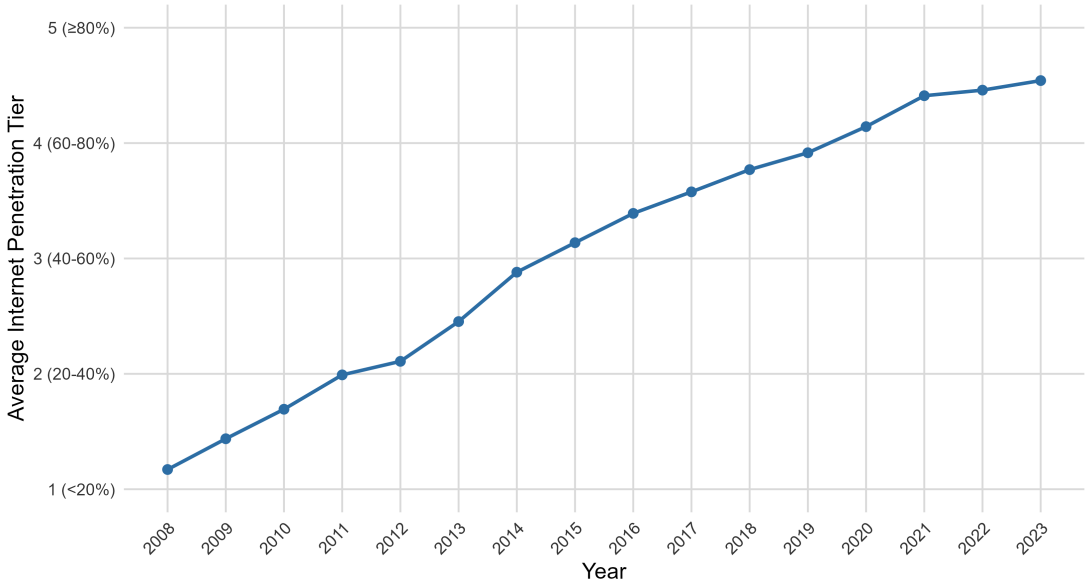
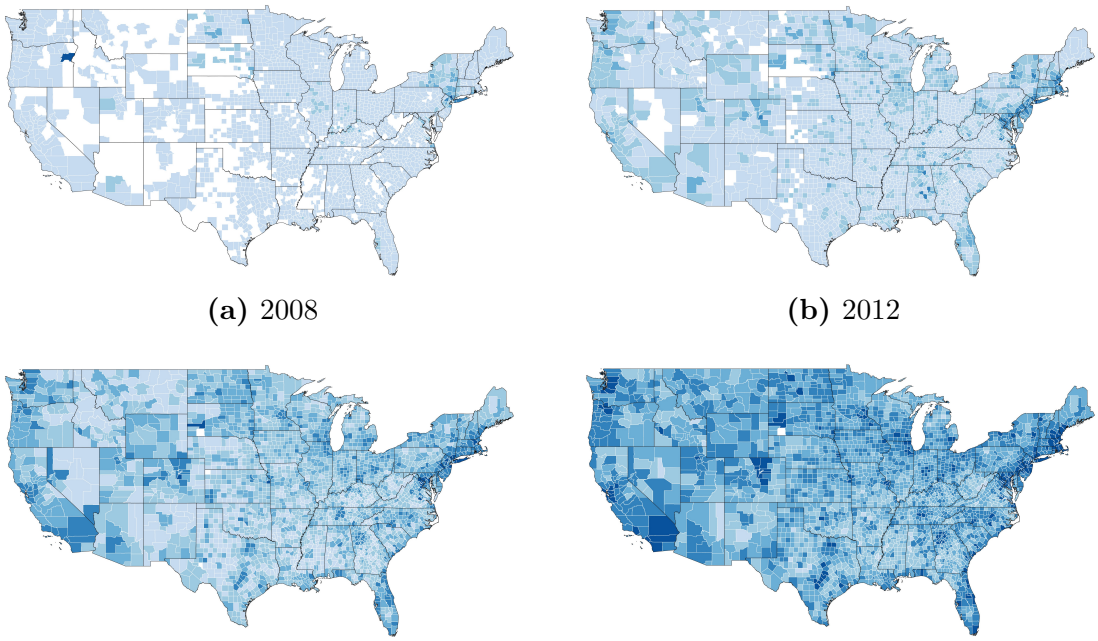


Figure 3.1: Evolution of High-Speed Internet Connectivity in the United States (2008–2023)

Notes: This figure plots the average weighted tier index of high-speed internet (10 Mbps or greater) connectivity across U.S. counties from 2008 to 2023, based on FCC broadband data. The tier index ranges from 0 to 5, where tier 1 indicates less than 20% of households with high-speed access and tier 5 indicates greater than 80% coverage.



(a) 2008 **(b)** 2012

(c) 2016 **(d)** 2020

Figure 3.2: Geographic Distribution of High-Speed Internet Access by Year

Notes: Each panel shows the county-level distribution of high-speed internet (10 Mbps or greater) access for the indicated year, based on FCC broadband data. Darker blue shading indicates a higher share of households with high-speed internet access; lighter shading indicates lower penetration. The maps illustrate the dramatic geographic expansion of high-speed connectivity from sparse rural coverage in 2008 to near-universal adoption by 2020, with the most rapid gains occurring between 2012 and 2016.

particularly pronounced among young adults aged 18-29, where prevalence rose from 0.7% to 4.3% over the same period. In addition, using the 2023 National Wellbeing Survey, London et al. (2025) show that among working-age adults (18-64 years), self-reported lifetime ADHD diagnosis prevalence has increased even more dramatically, rising from 4.25% in 2012 to an estimated 13.9% in 2023. This increase was evident across demographic groups, with particularly high rates among younger adults (20.9% among 18-29 year-olds).

This rise in ADHD prevalence appears to reflect substantial growth at the extensive margin of diagnosis, meaning that more individuals are being diagnosed with ADHD. However, increases in medication use need not map one-for-one into new diagnoses. They may also reflect growth at the extensive margin of treatment, if previously diagnosed but untreated individuals begin medication, and possibly changes at the intensive margin of treatment, such as more sustained or intensive use among treated patients. Evidence from national surveys and prescription data suggests that the increase has been driven importantly by expanded diagnosis among children (Danielson et al., 2024).

An important limitation of existing ADHD surveillance data is that publicly available surveys of diagnosis trends are conducted at the national or state level. The most comprehensive data sources, the National Health Interview Survey and National Survey of Children's Health do not provide county-level estimates. This geographic aggregation prevents researchers from linking ADHD diagnosis patterns to local variation in high-speed internet adoption, in contrast to the county-level analysis possible with pharmaceutical distribution databases. Consequently, while I can document the temporal correlation between rising broadband and video technology and ADHD diagnoses nationally, establishing more granular geographic correspondence requires using distribution data as a proxy for ADHD trends.

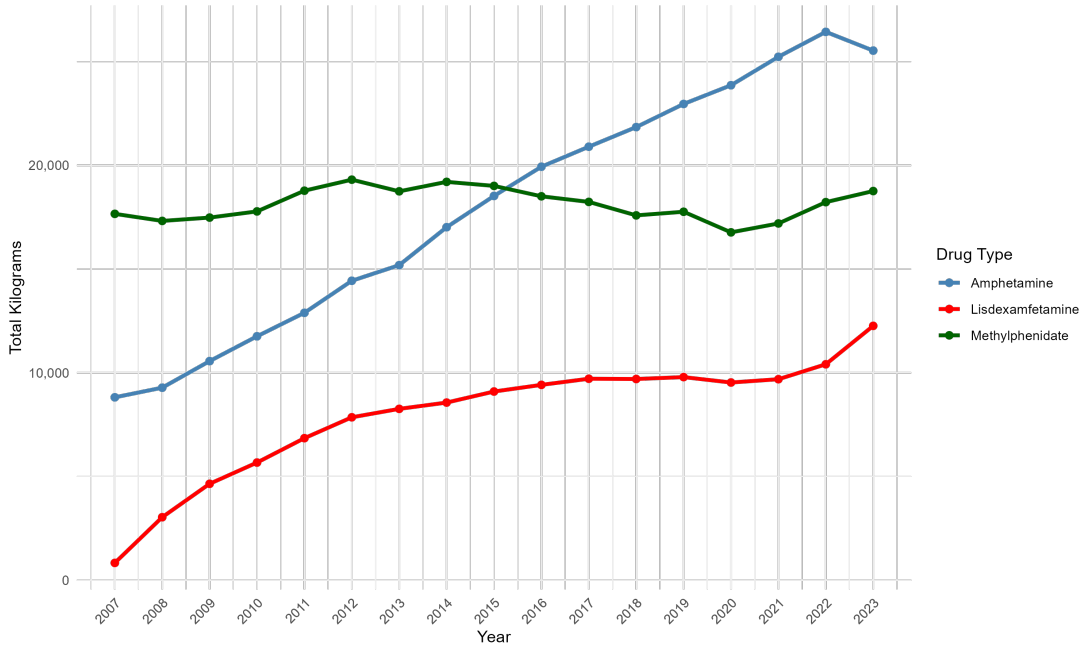
C. ADHD medication distribution. Figure 3.3 presents the national trends in ADHD medication distribution by drug type from 2007 to 2023. Amphetamine-based medications show the most dramatic increase, rising from approximately 8,000

kilograms in 2007 to over 25,000 kilograms by 2022, representing more than a threefold increase. Lisdexamfetamine exhibits the steepest growth trajectory, starting from near zero in 2007 and reaching approximately 12,000 kilograms by 2023, reflecting its introduction to the market and subsequent rapid adoption. Methylphenidate distribution remained relatively stable throughout the period, fluctuating around 18,000-19,000 kilograms annually. These trends coincide with the period of internet expansion, motivating the analysis that follows. Importantly, approximately half of diagnosed ADHD cases are treated with medication (Fletcher, 2014); therefore, the increase in real ADHD diagnoses due to internet use could be much higher than what medication distribution data alone suggests.

D. Special Education Enrollment. Although ADHD medication distribution increased significantly over the study period, the increase in special education identification rates was less pronounced. Figure 3.4 shows the evolution of Individualized Education Program (IEP) ratios in the United States from 2009 to 2019, where IEP ratio represents the number of students receiving special education services divided by total enrollment. The data reveals a modest increase in special education ratio, with the national average rising from approximately 14.0% in 2009 to 15.2% in 2019. This U-shaped pattern may reflect policy changes, diagnostic improvements, or increased awareness of special education needs during this period. Special education is a much broader umbrella category in which ADHD is included, and I do not have direct data on the number of ADHD-diagnosed students. However, Morrill (2018) indicates that a significant portion of special education needs stem from ADHD diagnoses. The analysis of special education outcomes covers a shorter time period than the ADHD medication analysis because publicly available data for special education ends in 2019.

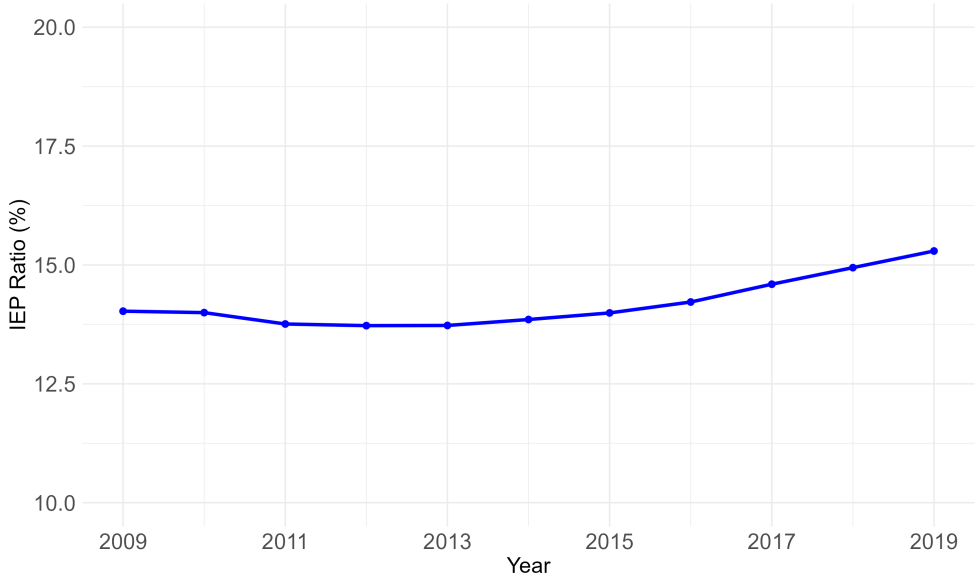
To summarize, ADHD diagnoses grew concurrent with the expansion of high-speed internet. These patterns are not evidence of a causal relationship on their own, so I turn to a two-way fixed effects approach to control for confounding trends and local characteristics that may be correlated with ADHD medications.

Figure 3.3: ADHD Medication Distribution Trends by Drug Type (2008–2023)



Notes: This figure plots national annual distribution of three ADHD medication classes (in kilograms) from 2007 to 2023, sourced from the DEA-ARCOS database. Amphetamine-based drugs (Adderall) are shown in one series, lisdexamfetamine (Vyvanse) in a second, and methylphenidate-based drugs (Ritalin/Concerta) in a third. Amphetamine-based medications show the most dramatic increase (from approximately 8,000 kg in 2007 to over 25,000 kg by 2022). Lisdexamfetamine grew from near zero following its market introduction. Methylphenidate remained relatively stable at 18,000–19,000 kg throughout the period. These trends coincide with the national expansion of high-speed broadband infrastructure.

Figure 3.4: Evolution of IEP Ratio in the United States (2009–2019)



Notes: This figure plots the national average IEP (Individualized Education Program) ratio from 2009 to 2019, defined as the number of students receiving special education services divided by total public school enrollment. Data are sourced from the Common Core of Data (CCD) maintained by the National Center for Education Statistics. The national average rose modestly from approximately 14.0% in 2009 to 15.2% in 2019. Special education is a broad category encompassing 13 disability types under IDEA, of which ADHD falls primarily under the "Other Health Impairment" classification. Data availability for special education ends in 2019, limiting the special education analysis to a shorter panel than the ADHD medication analysis.

3.2.3 Econometric Approach

I estimate a two-way fixed effects (TWFE) model to identify the effects of broadband expansion on ADHD and special education outcomes. The main specification is:

$$Y_{ct} = \alpha_c + \gamma_t + \beta \cdot HSI_{ct} + X'_{ct}\delta + \phi_1(f(t) \cdot \mathbf{Z}_c) + \epsilon_{ct} \tag{3.1}$$

where Y_{ct} represents the ADHD outcome variable for county c in year t. For medication demand, Y_{ct} is the natural logarithm of ADHD medication distribution from DEA-ARCOS data. For educational outcomes, Y_{ct} is the logarithm of special education enrollment from CCD. HSI_{ct} captures high-speed internet penetration ($\geq$ 10 Mbps) measured by the percent of connected households within each county. α_c and γ_t are county and year fixed effects, respectively, and X_{ct} is a vector of time-varying county controls including demographic and socioeconomic measures (percent in poverty, median age, median income, percent with a college degree, and population for ADHD medication regressions or total enrollment for special education regressions). The vector $\mathbf{Z}_c$ includes pre-adoption county-level characteristics measured at baseline (2007), and $f(t)$ is a linear time trend.

This specification addresses threats to identification beyond time-invariant county characteristics and overall time trends (captured by county and year fixed effects) by including interactions between baseline county characteristics and linear time trends. This approach mitigates confounding from within-county time-varying factors that correlate with both internet expansion and ADHD outcomes, such as shifts in parenting behaviors or healthcare practices in counties with different socioeconomic profiles.(Caldarulo et al., 2023; Carruthers & Wanamaker, 2013; Hoynes & Schanzenbach, 2009).

I employ two estimation approaches to identify the effect of high-speed internet on the demand for ADHD treatment drugs. First, I use high-speed internet penetration as a discrete treatment measure in tiers from 0 to 5, where tier 1 represents less than 20% of households with 10 Mbps connections, tier 2 represents between 20% and

40%, tier 3 between 40% and 60%, and so on. Second, I construct a binary treatment indicator for counties exceeding 40% high-speed penetration (tier 3 threshold).

The choice of tier 3 as the treatment threshold is motivated by theoretical and empirical considerations. First, when high-speed connection coverage approaches half of a county's housing units, the service has moved beyond early adopters to become broadly accessible to the general population, potentially creating network effects and widespread behavioral changes. Second, lower thresholds such as tier 1 may not provide sufficient time for internet connectivity to generate measurable effects on health outcomes. Infrastructure and usage patterns may still be nascent at less than 20% coverage, and attention disorders require time to develop and manifest clinically. Achieving tier 3 penetration typically implies that high-speed internet has been present in that county for at least one to two years. Third, higher thresholds such as 80-100% coverage (tier 5) are fairly uncommon until late in the 2008 - 2023 panel. Additionally, the primary effects of high-speed internet may have already been realized at that level of coverage, leaving little additional impact to detect.

To examine the dynamics of treatment effects, I estimate an event study specification using the binary treatment measure. Given the staggered rollout of broadband infrastructure across counties and potential heterogeneous treatment effects, I use difference-in-differences imputation as proposed by Borusyak et al. (2024). This technique imputes weighted counterfactual outcomes for treated units using not-yet-treated or never-treated control units.

The approach offers some advantages over traditional two-way fixed effects event study designs. First, it avoids the negative weighting problems that can arise when treatment effects are heterogeneous across counties or time (Goodman-Bacon, 2021; Sun & Abraham, 2021). Second, it provides more reliable inference by properly accounting for the uncertainty in the imputation process. Third, it allows for flexible modeling of treatment effect dynamics without imposing restrictive assumptions about parallel trends holding indefinitely into the future (Borusyak et al., 2024)

3.3 Results

Table 3.1 presents the baseline two-way fixed effects estimates examining the relationship between high-speed internet connectivity and total ADHD medication demand.Both specifications show a small, positive, and statistically significant relationship.

Model 1 employs a discrete-step measure of internet penetration based on the FCC's tiered classification system (0–5 scale), where each tier corresponds to a higher range of residential connections, from fewer than 200 connections per 1,000 households (tier 1) to more than 800 connections per 1,000 households (tier 5). A one-unit increase in this penetration scale—that is, movement to the next FCC tier, corresponding roughly to a 20 percentage point increase in household penetration—is associated with a 0.85 percent increase in total ADHD medication demand. This specification captures how ADHD medication demand changes as counties move through successive adoption thresholds.

The pattern is consistent with a progressively stronger relationship between broadband diffusion and ADHD medication demand at higher levels of penetration. One possible explanation is that gradual increases in internet access give residents more sustained exposure to digital leisure activities, allowing related habits to develop over time. Because the measure is defined at the county level, it may also capture peer and network effects: as broadband adoption expands within a county and more households engage in online gaming, streaming, or social media, others may be more likely to adopt similar behaviors. This type of local diffusion could reinforce the relationship between internet penetration and ADHD medication demand.

Model 2 employs a discrete-binary treatment approach, where the binary treatment indicator uses a threshold of tier 3 or higher, representing counties where at least 40% of housing units have access to high-speed internet connections (greater than 10Mbps). Figure 3.5 shows how the number of treated counties increased from 2008 to 2023. In 2008, most counties were classified as untreated (shown in gray),

with only a small fraction meeting the high-speed internet threshold. The transition accelerated markedly around 2012-2016, with the majority of counties crossing the treatment threshold by 2020. Model 2 shows treatment associated with a 1.14 percent increase in medication demand, significant at the 1% level.

These findings likely represent a lower bound of the true effect. Fletcher (2014) estimates that for every medicated ADHD case, at least one additional case goes undiagnosed or untreated, suggesting the actual impact on ADHD prevalence may be larger than medication data alone captures. I continue to explore potential mechanisms underlying these effects in more detail in Section 5.

Figure 3.6 displays dynamic treatment effects estimated using Borusyak et al. (2024). The pre-treatment coefficients are closely centered around zero. The post-treatment effects show a gradual upward pattern, with treatment effects reaching their peak around year 3 after crossing the threshold. Notably, the magnitude of the estimated effects using this estimator is approximately twice as large as those obtained from the standard TWFE estimation. This pattern suggests that the impact may take time to materialize, consistent with behavioral and developmental mechanisms that unfold under sustained exposure. Shifts in screen time and digital habits may take years to be reflected in ADHD-related treatment and medication demand.

Table 3.2 presents results for special education enrollment using IEP enrollment as the dependent variable. High-speed internet shows a less robust relationship with special education enrollment than with medication demand. The discrete-steps specification (0-5 scale) yields a positive and statistically significant coefficient: a 20% increase in connected households is associated with a 0.83% increase in special education enrollment. However, the binary treatment specification shows a larger but statistically insignificant effect. These more limited and imprecise results may be partially explained by the more constrained data availability for special education enrollment, which covers only the period from 2009-2019, compared to the longer time series available for ADHD medication data.

Table 3.1: Two-Way Fixed Effects: High-Speed Internet and ADHD Treatment Demand

	Log(Total ADHD Treatment Drugs Demand):	
	Discrete-Steps (0-5)	Discrete Treatment
	(1)	(2)
High-Speed Internet	0.0085***	0.0114***
	(0.0018)	(0.0026)
Observations	49,902	49,902
Controls	Yes	Yes
County & Year FE	Yes	Yes
Trend Interactions	Yes	Yes

Notes: ***p<0.01. Panel regressions with county and year fixed effects. Standard errors clustered at county level in parentheses. The dependent variable is log(ADHD drugs in grams) scaled by population. Trend interactions control for differential time trends based on 2007 baseline characteristics. Model 1: Multiple degrees of high-speed internet penetration (0–5 scale). Model 2: Binary treatment (1 if internet10m ≥ 3, indicating >40% of households in county i have 10 Mbps or greater internet connectivity).

Table 3.2: Two-Way Fixed Effects: High-Speed Internet and IEP Students

	Log(IEP Students)	
	Discrete Steps (0-5)	Discrete Binary
	(1)	(2)
High-Speed Internet	0.0083**	0.0110
	(0.0042)	(0.0076)
Observations	34,226	34,226
Controls	Yes	Yes
County & Year FE	Yes	Yes
Trend Interactions	Yes	Yes

Notes: *p<0.10, **p<0.05. Panel regression with county and year fixed effects. Standard errors clustered at county level in parentheses. The dependent variable is log(IEP students). Controls for total enrollment, demographics, and income included but not shown. Trend interactions control for differential time trends based on 2007 baseline characteristics. Model 1: Multiple degrees of internet access (0–5 scale). Model 2: Binary treatment (1 if internet10m ≥ 3). Sample covers years 2009–2019.

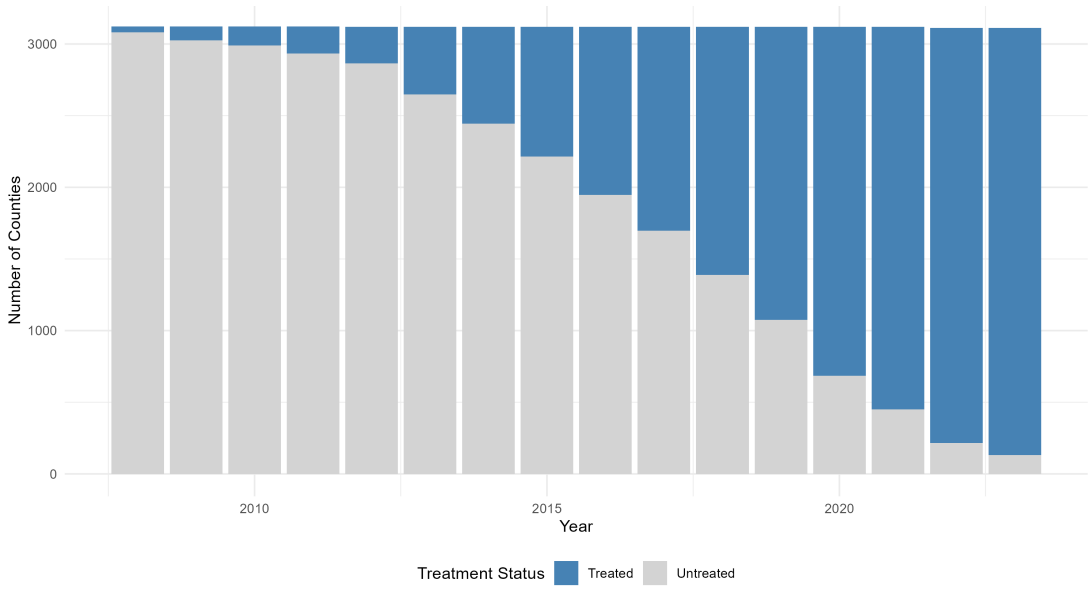


Figure 3.5: Number of Treated and Untreated Counties by Year

Notes: This figure shows the count of treated (counties with at least 40% of households having 10 Mbps or greater high-speed internet connectivity, i.e. tier ≥ 3) and untreated counties for each year from 2008 to 2023. Treated counties are shown in blue and untreated counties in gray. In 2008, only a small fraction of counties met the treatment threshold. The transition to treated status accelerated markedly between 2012 and 2016, with the majority of counties surpassing the threshold by 2020. This temporal variation in treatment adoption underpins the staggered difference-in-differences identification strategy.

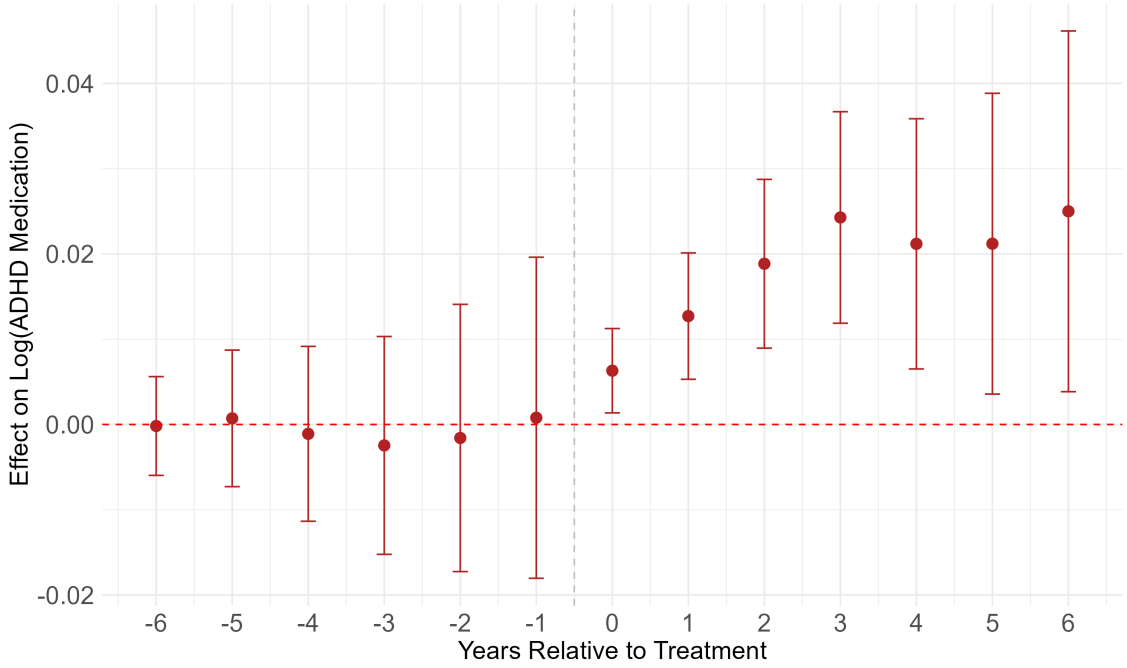


Figure 3.6: Event-Study Estimates – DID Imputation (Borusyak et al., 2024)

Notes: This figure displays event-study coefficients estimated using the Borusyak, Jaravel, and Spiess (2024) difference-in-differences imputation estimator. The dependent variable is log(ADHD prescription medication). The x-axis represents years relative to treatment adoption (counties crossing the tier ≥ 3 threshold), with period $t = -1$ normalized to zero as the reference period. Point estimates are shown with 95% confidence intervals. Pre-treatment coefficients are closely centered around zero, supporting the parallel trends assumption. Post-treatment effects show a clear upward trend peaking around year 3, with effect magnitudes approximately twice as large as those obtained from standard TWFE estimation, reflecting the correction for heterogeneous treatment effects in staggered adoption designs.

3.4 Alternative Specifications

To understand the robustness of my main findings and explore potential heterogeneity in the relationship between high-speed internet adoption and ADHD-related outcomes, I estimate several alternative specifications. First, I test whether the results are sensitive to alternative controls for differential time trends, including state-specific linear trends and year-specific interactions with baseline characteristics. Second, I examine the sensitivity of the findings to different high-speed internet penetration thresholds, demonstrating that the relationship holds across multiple definitions of treatment intensity. Third, I investigate heterogeneity by county population size to assess whether effects vary between rural and urban areas. Finally, I explore whether the overall increase in ADHD medication demand reflects broad changes across all therapeutic classes or is concentrated in specific drug types.

A. Alternative Time-Trend Controls. Following the approach in my main analysis, I use two alternative time trend specifications to control for within-county time-varying omitted variable bias. First, I add state-specific time trends to the baseline estimation, specifically, I estimate:

$$Y_{ct} = \alpha_c + \gamma_t + \beta \cdot HSI_{ct} + X'_{ct}\delta + \phi_1(f(t) \cdot \mathbf{Z}_c) + \theta_s \cdot f(t) + \epsilon_{ct} \tag{3.2}$$

where Y_{ct} represents ADHD medication demand for county c in year t, HSI_{ct} captures high-speed internet penetration ($\geq$ 10 Mbps), and α_c, γ_t, and X_{ct} are county fixed effects, year fixed effects, and time-varying county controls, respectively. The vector $\mathbf{Z}_c$ includes pre-adoption county-level characteristics measured at baseline (2007), $f(t)$ is a linear time trend, and θ_s is a vector of state-specific dummies that allows each state to have its own linear time trend.

Second, I modify the baseline specification by replacing the linear trend interaction $f(t)\cdot\mathbf{Z}_c$ with year-specific interactions, where baseline characteristics $\mathbf{Z}_c$ are interacted with a set of year dummy variables rather than a linear trend function.

Table 3.3: Alternative Time-Trend Controls

	Discrete-Steps (0-5)		Discrete Treatment	
	(1) Baseline	(2) + State Trends	(3) Baseline	(4) + State Trends
Panel A: State-Specific Linear Trends				
High-Speed Internet	0.0085***	0.0059***	0.0114***	0.0130***
	(0.0018)	(0.0016)	(0.0026)	(0.0022)
Panel B: Year Dummies Interacted with Baseline Characteristics				
High-Speed Internet	0.0060***	0.0035**	0.0057**	0.0083***
	(0.0019)	(0.0016)	(0.0028)	(0.0024)
County & Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
State-specific trends	No	Yes	No	Yes
Observations	49,902	49,886	49,902	49,886

Notes: *p<0.10, **p<0.05, ***p<0.01. Dependent variable is log(total ADHD drugs demand) at the county-year level. High-Speed Internet represents penetration of internet speeds $\geq$ 10 Mbps. In Panel A, discrete-steps (0–5) measures internet penetration on a 0–5 scale, while discrete treatment is an indicator equal to 1 if tier $\geq$ 3, indicating >40% of households in the county have 10 Mbps or greater internet connectivity. Panel B shows results when baseline county characteristics are interacted with year dummies instead of linear trends. All specifications include county and year fixed effects, plus time-varying county controls (median age, poverty rate, college degree rate, median income, log population). Standard errors clustered at the county level in parentheses. Positive and significant coefficients are consistent across all specifications, providing evidence that the main results are not driven by unobserved time-varying heterogeneity at the state or county level.

Table 3.3 shows that the effects remain robust across both specifications. In Panel A, when controlling for state-specific linear trends, the discrete-steps treatment effect decreases from 0.0085 to 0.0059, while the discrete treatment effect increases from 0.0114 to 0.0130. In Panel B, when using year dummies interacted with baseline characteristics instead of linear trends, the effects are smaller but still statistically significant: 0.0060 for the discrete specification and 0.0057 for the binary specification. The consistently positive and significant coefficients across specifications suggest the main results are not driven by unobserved time-varying heterogeneity at the state or county level.

B. Different High-Speed Internet Treatment Thresholds. I examine treatment effects across alternative penetration thresholds, as shown in Figure 3.7 and Table 3.4. As previously discussed, tier 3 serves as the primary treatment threshold because it represents the point where high-speed internet transitions from early adoption to broad accessibility. The positive association holds consistently across middle-range thresholds (tiers 2–4), with effect size increasing monotonically.

The extreme thresholds offer limited analytical value due to variation constraints. Tier 1 shows minimal variation early in the sample period, likely reflecting the nascent state of internet infrastructure, where effects have not yet materialized. Conversely, tier 5 demonstrates restricted variation throughout most of the observation period. The discrete-steps specification across all tiers confirms the overall positive relationship, supporting the choice of tier 3 as the optimal threshold that balances sufficient variation with meaningful treatment intensity.

C. Heterogeneity by County Size. I estimate the main specification on mutually exclusive subsamples defined by county population to examine whether effects vary by county size. Small counties are defined as those below the median population, medium counties fall between the 50th and 75th percentiles, large counties span the 75th to 90th percentiles, and very large counties comprise the top 10% by population.

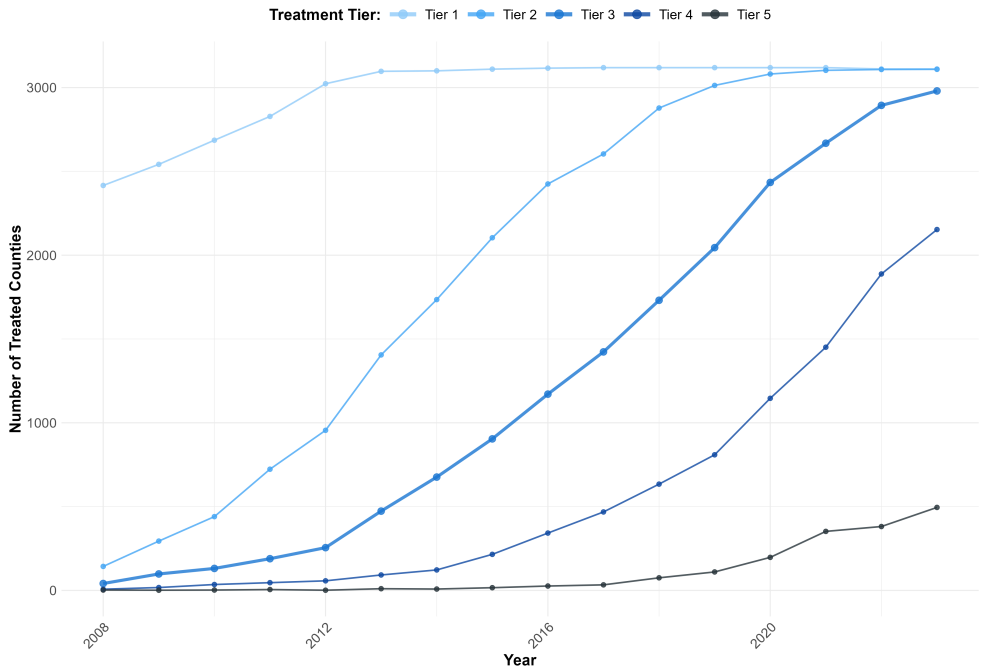


Figure 3.7: Counties Treatment Status per Tier-Threshold

Notes: Each line depicts the number of counties over time considered treated for the corresponding tier threshold. Tier 1 represents counties with 0–20% of households having 10+ Mbps connections, tier 2 represents 20–40%, tier 3 represents 40–60%, tier 4 represents 60–80%, and tier 5 represents counties with 80–100% of households having 10+ Mbps connections. Higher tiers indicate greater internet penetration ranges. The figure shows the temporal distribution of treatment adoption across different penetration levels, with tier 3 highlighted as the threshold used in the main specification. Tier 1 shows minimal variation early in the sample and tier 5 shows restricted variation throughout most of the observation period, limiting their analytical value relative to the middle-range thresholds.

The estimates, presented in Figure 3.8, indicate that the effect is largest in small counties, with a coefficient of about 0.015. The estimated effect declines monotonically with county size and is negative for very large counties. This pattern is consistent with geographic variation in ADHD prevalence documented in the medical literature, particularly across rural and less densely populated areas (Danielson et al., 2022, 2024). The negative coefficient for very large counties likely reflects limited identifying variation rather than a genuinely negative effect. As shown in Table 3.7, the estimated effect of broadband access is close to zero in the early period (2008–2013) and increases in later periods, implying that the main effects emerged after most large urban counties had already been treated*. As a result, there is little within-county variation left to identify the effect for this group.

The variation documented here may reflect greater vulnerability to technology-related disruptions in rural and smaller counties following the expansion of high-speed internet. One potential explanation is that youth in these areas may have had fewer recreational alternatives, leading to greater time allocation toward screen-based activities and gaming. I explore this mechanism in greater detail in the following section.

D. Heterogeneity by ADHD Drug Type. I investigate whether the increase in total ADHD medication demand was driven by a particular drug class or reflected broad changes across all medication types. To do so, I estimate separate models for: (i) amphetamine-based drugs (Adderall), (ii) lisdexamfetamine-based drugs (Vyvanse), and (iii) methylphenidate-based drugs (Ritalin, Concerta). Table 3.5 presents estimates for the discrete steps penetration measure (panel A) and binary treatment specification (panel B).

The results reveal heterogeneous effects across ADHD drug categories, reflecting distinct patient populations. Amphetamine-based medications (Adderall) show the largest response to internet penetration, with statistically significant coefficients of

*By 2014, 86.3% of very large counties had reached the tier 3 threshold, meaning that more than 40% of households already had high-speed internet access.

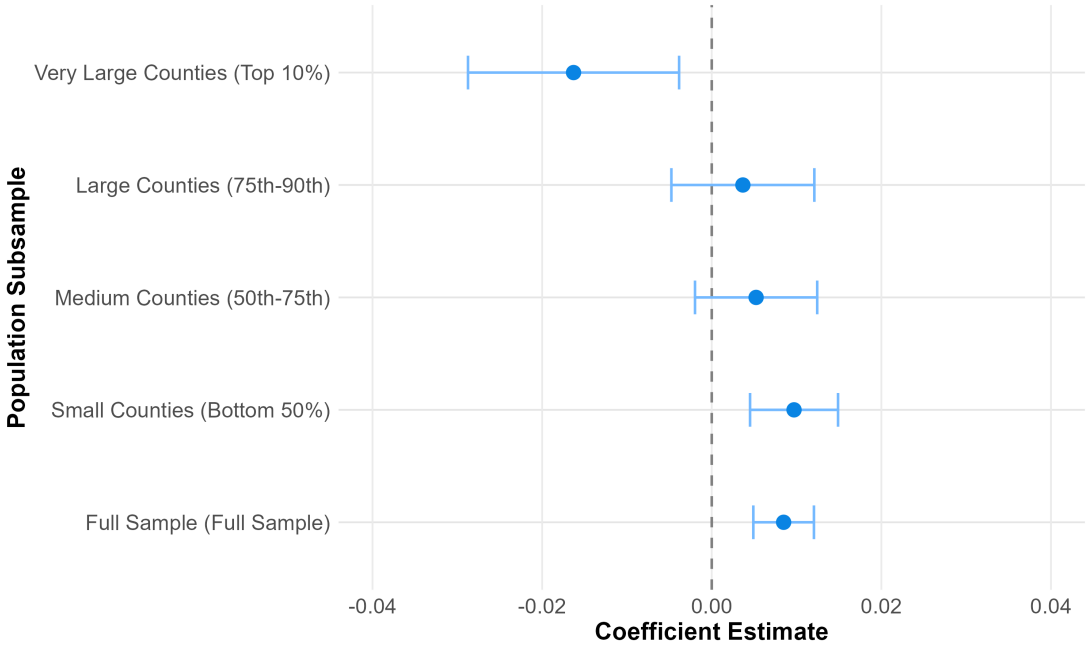


Figure 3.8: Heterogeneity Analysis by County Population Size

Notes: This figure presents TWFE coefficient estimates (with 95% confidence intervals) for the effect of high-speed internet on ADHD medication demand, estimated separately on four mutually exclusive county population subsamples. Small counties are defined as those below the median population; medium counties fall between the 50th and 75th percentiles; large counties span the 75th to 90th percentiles; and very large counties comprise the top 10% by population. The estimated effect is largest for small counties (centered around 0.015), this pattern is consistent with the medical literature documenting greater ADHD prevalence in rural and smaller communities, and may reflect fewer recreational alternatives for youth in these areas following high-speed internet expansion.

0.0165 under the discrete steps specification and 0.0245 under the binary treatment specification (approximately 1.7 and 2.5 percent increases, respectively). This is consistent with Adderall's widespread use among college students and young adults (Braghieri et al., 2022; IQVIA Institute for Human Data Science, 2024). Methylphenidate-based drugs (Ritalin) also exhibit significant positive responses with coefficients of 0.0106 (discrete) and 0.0150 (binary), representing approximately 1.1 and 1.5 percent increases, respectively. This aligns with methylphenidate's continued widespread use in traditional pediatric populations aged 6 years and older (Patel et al., 2024). In contrast, lisdexamfetamine-based medications (Vyvanse) show smaller, statistically insignificant coefficients of 0.0036 (discrete) and 0.0019 (binary) across both specifications.

3.5 Suggestive Evidence on Mechanisms

A. Leisure options and behavioral change. Following the leisure-work research design proposed by Aguiar et al. (2021), I use the American Time Use Survey (ATUS) to document a significant change in how young people allocate their time, shown in Figure 3.9. High-speed internet adoption increased from 2008 to 2022, closely paralleling increases in computer time among young males aged 15-20 years, whose weekly computing/gaming rose from approximately 7.5 hours in 2008 to 17.5 hours in 2023, an increase of more than 130%. In contrast, adult men aged 31-55 years exhibited stable computer usage throughout the period, hovering around 2-3.5 hours per week with no clear upward trend.

This descriptive pattern is consistent with evidence from Churchill and Johnson (2026), who find that a one-standard-deviation increase in state broadband coverage raised the likelihood that U.S. adolescents reported spending at least three hours online per day by 4.1 percentage points—an 11.6 percent increase relative to the sample mean.

Table 3.4: High-Speed Internet on ADHD Drug Demand by Different Thresholds

	Log(Total County Grams of ADHD Treatment Drugs)					
Treatment-Threshold	Tier ≥ 1 (1)	Tier ≥ 2 (2)	Tier ≥ 3 (3)	Tier ≥ 4 (4)	Tier ≥ 5 (5)	Discrete-Steps (0-5) (6)
High-Speed Internet	-0.0141^{*} (0.0062)	0.0067^{**} (0.0025)	0.0114^{***} (0.0026)	0.0184^{***} (0.0031)	-0.0218^{***} (0.0065)	0.0085^{***} (0.0018)
County & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls & Trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,902	49,902	49,902	49,902	49,902	49,902

Notes: *p<0.05, **p<0.01, ***p<0.001. Panel regressions with county and year fixed effects. Trend interactions control for differential time trends based on 2007 baseline characteristics. Standard errors clustered at county level in parentheses. Binary treatment models: 1 if internet10m $\geq$ threshold, 0 otherwise. Continuous model: internet penetration on 0–5 scale. All models include controls for median age, poverty rate, college education rate, median income, log population, and their interactions with 2007 baseline characteristics. The positive association between high-speed internet and ADHD medication demand is consistently observed across middle-range thresholds (tiers 2, 3, and 4), while extreme thresholds (tiers 1 and 5) show limited variation and less reliable estimates.

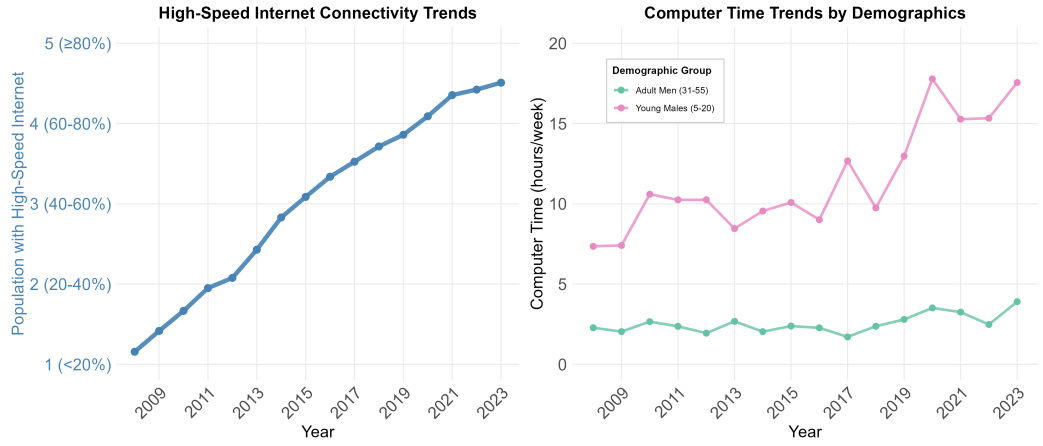


Figure 3.9: High-Speed Internet Adoption and Computer Time by Age Group

Notes: The left panel plots the weighted average tier index of high-speed internet (10 Mbps or greater) connectivity across U.S. counties from 2008 to 2023, based on FCC broadband data. Tier 1 indicates less than 20% penetration, tier 2 indicates 20–40%, tier 3 indicates 40–60%, tier 4 indicates 60–80%, and tier 5 indicates greater than 80% penetration. The right panel plots weekly average computer time (in hours) by demographic group from the American Time Use Survey (ATUS): young males aged 15–20 years (pink line) and adult men aged 31–55 years (green line). Weekly computing and gaming among young males rose from approximately 7.5 hours in 2008 to 17.5 hours in 2023 (an increase exceeding 130%), while adult men's usage remained relatively stable at 2–3.5 hours per week throughout the period.

Table 3.5: High-Speed Internet and ADHD Medications by Drug Type

	Log(County-level ADHD Drug Grams)		
	Amphetamine (Adderall)	Lisdexamfetamine (Vyvanse)	Methylphenidate (Ritalin)
	(1)	(2)	(3)
Panel A: Discrete Steps Specification			
High-Speed Internet Penetration (0-5)	0.0165***	0.0036	0.0106***
	(0.0025)	(0.0042)	(0.0020)
Panel B: Binary Treatment Specification			
High-Speed Internet Prevalence	0.0245***	0.0019	0.0150***
	(0.0036)	(0.0063)	(0.0027)
County & Year FE	Yes	Yes	Yes
Controls & Time-Trends	Yes	Yes	Yes
Observations	49,902	49,902	49,902

Notes: ***p<0.01, **p<0.05, *p<0.1. Panel regressions with county and year fixed effects. Standard errors clustered at county level in parentheses. Panel A uses continuous internet penetration measure (0–5 scale). Panel B uses binary treatment (Treatment = 1 if internet tier $\geq$ 3, which means more than 40% of households have high-speed internet connectivity). Controls for demographics and time trends with baseline characteristics included but not shown. Amphetamine-based medications (Adderall) show the largest response (coefficients of 0.0165 and 0.0245 in Panels A and B, respectively), consistent with Adderall's widespread use among college students and young adults. Methylphenidate (Ritalin) also shows significant positive effects. Lisdexamfetamine (Vyvanse) shows smaller, statistically insignificant coefficients across both specifications.

I formalize this relationship in Table 3.6 using a TWFE regression of high-speed internet connectivity on computer time across demographic groups. Because the ATUS does not provide county-level identifiers, I aggregate the data to the state level and estimate specifications with state and year fixed effects and demographic controls. The coefficients on high-speed internet connectivity are positive across all groups, and the point estimate for young males is substantially larger than those for adult men and the full sample. Although the estimates are imprecise because the analysis is conducted at the state-year level, the pattern is consistent with high-speed internet primarily reshaping how young people spend their time.

This shift in recreational habits may have produced neurological and behavioral changes that could reduce capacity for sustained attention and concentration. Intensive exposure to the rapid stimuli of online gaming, streaming, and social media may alter neural pathways related to attention regulation and impulse control (Swing et al., 2010; Twenge et al., 2018). Children and adolescents are particularly susceptible to such influences, as their brains are still developing (Colombo & Failache, 2023). Consequently, the rise in ADHD diagnoses and medication demand may partly reflect technology-driven changes in daily habits.

Building on this framework, I examine whether the effects of high-speed internet access vary across different time periods, recognizing that online gaming culture requires time to develop and depends on peer participation for sustained engagement. High-speed internet access may have had progressively stronger effects on ADHD treatment demand in later periods, as gaming communities matured and online activities became more prevalent among youth. As Figure 3.9 illustrates, computing time among young people surged notably after 2016, suggesting that the behavioral impacts of internet access intensified as digital entertainment became more sophisticated and engaging.

To test this temporal variation hypothesis, I estimate the following model that interacts high-speed internet access with distinct time periods:

$$\begin{aligned}\log(Y_{ct}) = \alpha_c + \gamma_t + \beta_1 HSI_{ct} + \beta_2 HSI_{ct} \times \text{Period2}_t \\ + \beta_3 HSI_{ct} \times \text{Period3}_t + \beta_4 \text{Period2}_t + \beta_5 \text{Period3}_t \\ + \boldsymbol{X}'_{ct}\boldsymbol{\delta} + \boldsymbol{Z}'_{c,2007} \times t\boldsymbol{\phi} + \epsilon_{ct}\end{aligned} \tag{3.3}$$

where Y_{ct} represents total ADHD treatment drug demand in county c during year t, HSI_{ct} is the high-speed internet access indicator, Period2_t and Period3_t are dummy variables for the periods 2016–2019 and 2020–2023 respectively (with 2008–2015 as the reference period), $\boldsymbol{X}'_{ct}$ includes time-varying county controls, $\boldsymbol{Z}'_{c,2007} \times t$ represents trend interactions based on 2007 baseline characteristics, α_c and γ_t denote county and year fixed effects, and ϵ_{ct} is the error term. The coefficients β_2 and β_3 capture the differential effects of high-speed internet access in the later periods relative to the baseline period. Importantly, Period 3 (2020–2023) coincides with the COVID-19 public health emergency, during which the DEA suspended the in-person evaluation requirements of the Ryan Haight Act of 2008, enabling telemedicine prescribing of controlled substances without a prior in-person visit. This policy shift may have independently contributed to rising ADHD medication demand during this period, making it difficult to fully attribute the estimated effects of high-speed internet access solely to behavioral channels.

The results in Table 3.7 offer support for the temporal variation hypothesis. The baseline coefficient for Period 1 (2008–2015) is positive and statistically significant in Model 1, suggesting that internet access was already associated with higher ADHD treatment demand during the early broadband expansion period. The interaction terms in Period 2 (2016–2019) show larger and more significant effects across both models, consistent with the hypothesis that behavioral impacts of high-speed internet strengthened as online gaming ecosystems matured and digital entertainment became more immersive.

Interpreting the Period 3 (2020–2023) interactions requires caution. While the estimated effects remain significant — and notably larger in the binary treatment specification — this period coincides with the DEA's COVID-era suspension of the Ryan Haight Act's in-person prescribing requirements, which substantially eased access to controlled substance prescriptions via telemedicine. The stronger interaction observed in Model 2 during Period 3 may therefore reflect a combination of continued behavioral effects and easier drug access policy changes. The present research design does not allow for a clean separation of these two channels. Nevertheless, significant effects are observed consistently across all three periods — including Period 1 and Period 2, which predate any telehealth policy changes — suggesting that the behavioral channel is not merely an artifact of easier prescribing access, and that sustained exposure to online gaming and digital entertainment carries independent explanatory power.

B. Awareness/Information channel. An alternative explanation for rising ADHD diagnoses is increased awareness and information access rather than genuine behavioral change. By 2008, major search engines and basic internet connectivity were widely available, enabling parents, teachers, and healthcare providers to research ADHD symptoms and treatment options. During this period, ADHD diagnoses had already increased substantially among children, coinciding with growing knowledge about the condition.

To examine whether the relationship between internet access and ADHD medication demand varies by connection speed, I estimate the same regression specifications using low-speed internet (200Kbps instead of 10Mbps) penetration as the treatment variable. Low-speed internet connectivity serves as a proxy for information access provided by the internet to parents and healthcare providers, as these connections already allowed users to look up medical information and research ADHD symptoms. This approach addresses whether basic internet connectivity and information access alone drive changes in ADHD medication demand, or whether the relationship depends on high-speed connections that enable different types of digital engagement.

On Table 3.8/panel A, I estimate the effect of low-speed internet penetration on ADHD medication demand (instead of high-speed internet), while on panel B, I include both high-speed and low-speed internet variables simultaneously. High-speed internet shows a larger and more precisely estimated effect on ADHD medication demand, with coefficients (0.0077 and 0.0113) exceeding those of low-speed internet across both specifications. While both channels likely played some role, the results suggest that behavioral change weighs more heavily than information access in driving medication demand.

3.6 Discussion

I investigate the relationship between high-speed internet expansion and ADHD-related outcomes in the United States from 2008 to 2023, using comprehensive county-level data on pharmaceutical distribution, special education enrollment, and broadband connectivity. This research provides the first large-scale quasi-experimental analysis of how high-speed internet expansion may affect attention-related disorders across diverse geographic contexts.

Using two-way fixed effects models with county and year fixed effects, I find that when counties surpassed 40% high-speed internet adoption, total demand for ADHD medications increased by 1.14%. A multiple-discrete measure of internet penetration yields similar conclusions, with each tier increase associated with a 0.85 percent increase in medication demand. Across the full set of robustness checks, the estimated coefficients remain positive and statistically consistent in direction, ranging from 0.35% to 1.84% depending on the specification and treatment threshold considered. A difference-in-differences imputation estimator (Borusyak et al., 2024) that accounts for heterogeneous treatment effects yields magnitudes roughly twice as large.

When applying the same framework to special education enrollment, I observe a positive and statistically significant coefficient of 0.83% under the discrete-steps

specification, while the binary treatment yields a larger but insignificant effect of 1.10%. These more limited results likely reflect the broader scope of special education beyond ADHD, shorter data availability (2009-2019), and institutional factors, including state financial incentives.

Effects are not driven by substitution between drug classes but reflect broad increases across therapeutic categories, with amphetamine-based medications (Adderall) showing the largest response (2.45%) and methylphenidate-based drugs exhibiting a 1.50% increase. The pronounced effect on Adderall aligns with its widespread adoption among college students and young adults, consistent with National Health Interview Survey data showing youth experienced the largest increases in new ADHD diagnoses during this period. Heterogeneity analysis reveals more pronounced effects in smaller counties.

Data from the American Time Use Survey reveals substantial increases in computer time among young males aged 15-20 years during the high-speed internet expansion period, with weekly computing and gaming rising from approximately 7.5 hours in 2008 to 17.5 hours in 2023—an increase exceeding 130%. TWFE regressions at the state level show that high-speed internet connectivity correlates positively with computer time across all demographic groups, with point estimates for young males larger in magnitude than those for other groups. Given the imprecision of these estimates, the effects appear to operate primarily through changes in how young people spend their time rather than through mechanisms affecting all age groups equally.

While increased awareness and information access may partly explain the findings, several factors point toward behavioral and neurological changes as the more important contributors. The technological specificity of the observed effects—with high-speed connectivity showing distinct impacts compared to basic internet access—suggests different patterns of use enabled by bandwidth-intensive applications. Supporting evidence from Uruguay (Colombo & Failache, 2023) and Spain (Arenas-Arroyo et al., 2025) indicates that high-speed internet exposure can adversely affect

child development and adolescent mental health when parental supervision is limited. The gradual emergence of treatment effects over time, combined with dramatic increases in youth screen time, suggests that sustained behavioral changes drive a substantial portion of the documented increase in ADHD-related healthcare demand. Moreover, even if awareness plays some role, there is little reason to expect this channel to systematically correlate with the timing and geography of broadband adoption in ways that would confound identification.

These findings carry important policy implications for ongoing federal broadband expansion initiatives. While high-speed internet generates substantial benefits for education, economic opportunity, and social connectivity (Caldarulo et al., 2023; Dettling et al., 2018; Grimes & Townsend, 2018), this research highlights developmental costs that merit consideration in policy design. The Connect America Fund has allocated over $10 billion to expand broadband to underserved rural areas—precisely the smaller counties where this study documents more pronounced increases in ADHD-related healthcare demand. Future broadband expansion policies might pair infrastructure investments with targeted interventions — digital literacy programs, parental guidance resources, and community activities — particularly in smaller counties.

The study acknowledges some limitations that suggest directions for future research. County-level aggregation may mask important household and individual heterogeneity in responses to internet connectivity. Administrative pharmaceutical data, while comprehensive, cannot capture undiagnosed cases or non-pharmaceutical treatment approaches. The observational research design, despite robust identification strategies, cannot definitively establish causal mechanisms. Individual-level longitudinal data could further illuminate these relationships and inform more targeted policy responses.

Table 3.6: TWFE – High-Speed Internet and Computer Time Use

	Log(Avg Computer Time)		
	Young Males (15-20)	Adult Men (31-55)	All
	(1)	(2)	(3)
High-Speed Internet Connectivity (0/5)	0.160	0.033	0.046
	(0.202)	(0.108)	(0.060)
State & Year FE	Yes	Yes	Yes
Controls & Trends	Yes	Yes	Yes
Observations	699	810	812

Notes: Panel regressions with state and year fixed effects. The dependent variable is log(average weekly computer time in hours) for the indicated demographic group, computed from the American Time Use Survey (ATUS). Computer time and total population are log-transformed. Trend interactions control for differential time trends by 2007 baseline characteristics. Standard errors clustered at the state level are reported in parentheses. Controls include Median Age, Percent in Poverty, Percent College Degree, Median Income, and Log(Total Population). The coefficient on high-speed internet connectivity is positive in all columns and largest for young males (15–20), though the estimates are imprecise because ATUS aggregation limits the analysis to a relatively small number of state-year observations.

Table 3.7: Time Period Sub Groups – Heterogeneous Treatment Effects

	Log(Total County ADHD Treatment Drugs Demand):	
	Internet Penetration	Internet Treatment
	(1)	(2)
High-Speed Internet	0.0044**	0.0054
	(0.0021)	(0.0034)
High-Speed Internet × Period 2 (2016–2019)	0.0082***	0.0070**
	(0.0019)	(0.0032)
High-Speed Internet × Period 3 (2020–2023)	0.0067**	0.0137**
	(0.0033)	(0.0063)
Observations	49,902	49,902
Controls & Trends	Yes	Yes
County & Year FE	Yes	Yes

Notes: **p<0.05, ***p<0.01. Panel regression with county and year fixed effects using the full sample. Standard errors clustered at county level in parentheses. Period 1 (reference): 2008–2015; Period 2: 2016–2019; Period 3: 2020–2023. Model 1 uses continuous internet penetration (0–5 scale). Model 2 uses a binary internet treatment indicator. Time-trend interactions control for differential trends based on 2007 baseline characteristics. The baseline coefficient reflects the early high-speed internet period (2008–2015). Period 2 (2016–2019) captures the pre-COVID telehealth era, during which the Ryan Haight Act of 2008 required an in-person medical evaluation prior to any controlled substance prescription, including ADHD medications; internet access nonetheless shows a significant positive association with demand, consistent with growing behavioral impacts as online gaming ecosystems matured and digital entertainment became more immersive. Period 3 (2020–2023) coincides with the COVID-19 public health emergency and the DEA's suspension of Ryan Haight Act in-person requirements, enabling telemedicine prescribing of controlled substances without a prior in-person visit. The stronger and more significant interaction in Model 2 during Period 3 may therefore reflect a combination of continued behavioral effects of high-speed internet access and expanded prescribing access via telehealth, making it difficult to fully disentangle easier drug access policy changes from the behavioral shifts' effects in this period.

Table 3.8: Low-Speed Internet Penetration and ADHD Treatment Drug Demand

	Log(Total County ADHD Treatment Drugs Demand)	
	Discrete-Steps (0-5)	Discrete-Binary
	(1)	(2)
Panel A: Low-Speed Internet Only		
Low-Speed Internet	0.0060***	-0.0027
	(0.0019)	(0.0033)
Panel B: Both High-Speed and Low-Speed Internet		
High-Speed Internet	0.0077***	0.0113***
	(0.0018)	(0.0026)
Low-Speed Internet	0.0042**	-0.0019
	(0.0018)	(0.0033)
Observations	49,902	49,902
County & Year FE	Yes	Yes
Time-Trend & Controls	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. Panel regressions with county and year fixed effects. Trend interactions control for differential time trends based on 2007 baseline characteristics. Standard errors clustered at county level in parentheses. Column 1 shows continuous internet penetration (0–5 scale). Column 2 shows binary treatment indicators (1 if $\geq$ 3). High-speed internet refers to 10M+ connections; low-speed internet refers to 200k+ connections. All regressions include controls for median age, percent in poverty, percent college degree, median income, and log population. Panel A estimates the effect of low-speed internet penetration alone, as a proxy for general information access. Panel B includes both high-speed and low-speed variables simultaneously. Across both panels, high-speed internet coefficients consistently exceed those of low-speed internet in magnitude and statistical precision, suggesting that the behavioral change channel (enabled by bandwidth-intensive activities) carries greater weight than the information access channel in driving the observed increases in ADHD medication demand.

References

Aguiar, M., Bils, M., Charles, K. K., & Hurst, E. (2021). Leisure luxuries and the labor supply of young men. *Journal of Political Economy*, *129*(2), 337–382. https://doi.org/10.1086/711916

Akinbami, L. J., Liu, X., Pastor, P. N., & Reuben, C. A. (2011). Attention deficit hyperactivity disorder among children aged 5–17 years in the United States, 1998–2009. *NCHS Data Brief*, (70). https://www.cdc.gov/nchs/products/databriefs/db70.htm

Arenas-Arroyo, E., Fernandez-Kranz, D., & Nollenberger, N. (2025). High speed internet and the widening gender gap in adolescent mental health: Evidence from Spanish hospital records [Forthcoming]. *Journal of Health Economics*. https://doi.org/10.1016/j.jhealeco.2025.103014

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(5), 3253–3285. https://doi.org/10.1093/restud/rdae007

Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American Economic Review*, *112*(11), 3660–3693. https://doi.org/10.1257/aer.20211218

Caetano, C., Caetano, G. S., Mazetto, D., & Skira, M. (2025). *Is video watching bad for kids? The effect of video watching on children's skills* (tech. rep.). National Bureau of Economic Research.

Caldarulo, M., Mossberger, K., & Howell, A. (2023). Community-wide broadband adoption and student academic achievement. *Telecommunications Policy*, *47*, 102445. https://doi.org/10.1016/j.telpol.2022.102445

Carruthers, C. K., & Wanamaker, M. H. (2013). Closing the gap? The effect of private philanthropy on the provision of African-American schooling in the US south. *Journal of Public Economics*, *101*, 53–67.

CHADD. (2025). *Individuals with disabilities education act*. Children and Adults with Attention-Deficit/Hyperactivity Disorder. Retrieved January 24, 2025, from https://chadd.org/for-parents/individuals-with-disabilities-education-act/

Chung, W., Jiang, S.-F., Paksarian, D., Nikolaidis, A., Castellanos, F. X., Merikangas, K. R., & Milham, M. P. (2019). Trends in the prevalence and incidence of attention-deficit/Hyperactivity disorder among adults and children of different racial and ethnic groups. *JAMA Network Open*, *2*(11), e1914344. https://doi.org/10.1001/jamanetworkopen.2019.14344

Churchill, B. F., & Johnson, K. R. (2026). *Broadband internet access and adolescent mental health in the US* (tech. rep.). National Bureau of Economic Research.

Colombo, K., & Failache, E. (2023, November). *Exposure to high-speed internet and early childhood development: Evidence from a countrywide program* (Working Paper No. 2023-01) (First version: January 19, 2023. This version: November 1, 2023). GEAR - Graduate Program in Applied Economic Research, Universitat Autònoma de Barcelona.

Danielson, M. L., Claussen, A. H., Bitsko, R. H., Holbrook, J. R., Charania, S. N., Haynie, D., Rohan, J. M., Rast, J. E., & Kaminski, J. W. (2024). ADHD prevalence among U.S. children and adolescents in 2022: Diagnosis, severity, co-occurring disorders, and treatment [Published online May 22, 2024]. *Journal of Clinical Child & Adolescent Psychology*. https://doi.org/10.1080/15374416.2024.2335625

Danielson, M. L., Holbrook, J. R., Bitsko, R. H., Newsome, K., Charania, S. N., McCord, R. F., Kogan, M. D., & Blumberg, S. J. (2022). State-level estimates of the prevalence of parent-reported ADHD diagnosis and treatment among US children and adolescents, 2016 to 2019. *Journal of Attention Disorders*, *26*(13), 1685–1697.

Dettling, L. J., Goodman, S., & Smith, J. (2018). Every little bit counts: The impact of high-speed internet on the transition to college. *The Review of Economics and Statistics*, *100*(2), 260–273. https://doi.org/10.1162/rest_a_00712

Donati, D., Durante, R., Sobbrio, F., & Zejcirovic, D. (2025). Lost in the net? Broadband internet and youth mental health. *Journal of Health Economics*, 103017.

Fletcher, J. M. (2014). The effects of childhood ADHD on adult labor market outcomes. *Health Economics*, *23*(2), 159–181. https://doi.org/10.1002/hec.2907

Geraci, A., Nardotto, M., Reggiani, T., & Sabatini, F. (2022). Broadband internet and social capital. *Journal of Public Economics*, *206*, 104578.

Golin, M. (2022). The effect of broadband internet on the gender gap in mental health: Evidence from Germany. *Health Economics*, *31*(S2), 6–21. https://doi.org/10.1002/hec.4570

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Grimes, A., & Townsend, W. (2018). Effects of (ultra-fast) fibre broadband on student achievement. *Information Economics and Policy*, *44*, 8–15. https://doi.org/10.1016/j.infoecopol.2018.06.001

Hoynes, H. W., & Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, *1*(4), 109–139. https://doi.org/10.1257/app.1.4.109

IQVIA Institute for Human Data Science. (2024). *Stimulant prescription trends in the United States from 2012–2023*. U.S. Drug Enforcement Administration. https://www.deadiversion.usdoj.gov/pubs/docs/IQVIA-Report-on-Stimulant-Trends-2024.pdf

Jain, R., & Stemper, S. (2025). 3G internet and human development. *Working Paper*.

London, A. S., Monnat, S. M., & Gutin, I. (2025). Self-reported ADHD diagnosis status among working-age adults in the United States: Evidence from the 2023 National Wellbeing Survey. *Journal of Attention Disorders*, *29*(6), 399–410. https://doi.org/10.1177/10870547251319861

Morrill, M. S. (2018). Special education financing and ADHD medications: A bitter pill to swallow. *Journal of Policy Analysis and Management*, *37*(2), 384–402. https://doi.org/10.1002/pam.22055

National Center for Education Statistics. (2023). *Students with disabilities*. U.S. Department of Education. Retrieved October 13, 2023, from https://nces.ed.gov/programs/coe/indicator/cgg/students-with-disabilities

Patel, P., et al. (2024). Methylphenidate [Updated October 29, 2024]. In *Statpearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK482451/

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Swing, E. L., Gentile, D. A., Anderson, C. A., & Walsh, D. A. (2010). Television and video game exposure and the development of attention problems. *Pediatrics*, *126*(2), 214–221. https://doi.org/10.1542/peds.2009-1508

Twenge, J. M., Martin, G. N., & Campbell, W. K. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. *Emotion*, *18*(6), 765.

Zuo, G. W. (2021). Wired and hired: Employment effects of subsidized broadband internet for low-income Americans. *American Economic Journal: Economic Policy*, *13*(3), 447–482.

Appendix A

This appendix presents a robustness check for the main results. Baker County, Oregon (FIPS 41001) was found to contain anomalous high-speed internet penetration values in 2008, likely reflecting a data recording error. Table A1 replicates the baseline specifications from Table 3.1 after excluding this county.

Table A1: Two-Way Fixed Effects: High-Speed Internet and ADHD Treatment Demand (Robustness Check: Excluding Baker County, Oregon)

	Log(Total ADHD Treatment Drugs Demand):	
	Discrete-Steps (0-5)	Discrete Treatment
	(1)	(2)
High-Speed Internet	0.0085***	0.0114***
	(0.0018)	(0.0026)
Observations	49,886	49,886
Controls	Yes	Yes
County & Year FE	Yes	Yes
Trend Interactions	Yes	Yes

Notes: ***p<0.01. Panel regressions with county and year fixed effects. Standard errors clustered at county level in parentheses. The dependent variable is log(ADHD drugs in grams) scaled by population. Trend interactions control for differential time trends based on 2007 baseline characteristics. Model 1: Multiple degrees of high-speed internet penetration (0–5 scale). Model 2: Binary treatment (1 if internet10m $\geq$ 3, indicating >40% of households in county i have 10 Mbps or greater internet connectivity). Baker County, Oregon (FIPS 41001) excluded due to anomalous internet10m values detected in 2008.

Chapter 4

Sales Tax Enforcement and Business Activity: Evidence from Economic Nexus Laws

Abstract

The *South Dakota v. Wayfair* (2018) decision authorized states to impose sales tax obligations on remote sellers based on economic presence rather than physical location, prompting nearly universal adoption of economic nexus thresholds across U.S. states. While most states converged toward a $100,000 annual sales threshold, meaningful variation persists in how binding these thresholds are relative to state market size. This paper examines how economic nexus threshold exposure—measured as the threshold amount normalized by state population—affects business activity using a state-level panel from 2013 to 2023. Exploiting cross-state variation in threshold tightness following *Wayfair*, we estimate two-way fixed effects models with a rich set of controls and fixed effects. We find small but statistically significant negative effects of higher threshold exposure on the number of firms, establishments, and employment, with effects concentrated among very small firms (fewer than 20 employees). These results suggest that states where the nexus threshold is less binding relative to market size experience modest competitive disadvantages for physical retailers, as more online sellers remain exempt from tax collection obligations, widening tax-induced price differentials. The findings challenge the assumption that higher thresholds are neutral for economic activity and underscore that threshold exposure—not merely threshold level—is the relevant policy parameter for ensuring competitive equity between online and brick-and-mortar commerce.

4.1 E-commerce Taxation and Business Activity

Until 2018, U.S. sales tax law required retailers to have a physical presence in a state before being obligated to collect and remit sales taxes — a standard established by the Supreme Court in *Quill Corp. v. North Dakota* (1992). As e-commerce expanded rapidly through the 2000s and 2010s, this physical presence rule created a significant competitive asymmetry: remote online sellers could offer goods free of sales tax while brick-and-mortar retailers faced rates that sometimes reached 10%, disadvantaging local businesses and eroding state tax bases (Beem & Bruce, 2021; Bruce & Fox, 2000). Conservative estimates suggest these untaxed electronic transactions produced cumulative revenue shortfalls of $52 billion for state and local governments between 2007 and 2012 (Beem & Bruce, 2021). The Supreme Court's decision in *South Dakota v. Wayfair* (2018) transformed this landscape by overturning *Quill* and authorizing states to impose tax obligations based on *economic presence* alone, empowering states to require remote sellers to collect sales taxes even without physical operations within their borders.

In the wake of *Wayfair*, nearly every U.S. state adopted an economic nexus policy requiring remote sellers to remit sales tax once certain thresholds were surpassed. These thresholds—initially varied—converged over time toward a common baseline of $100,000 in sales and/or 200 transactions, though a few states, including New York, California, and Texas, opted for substantially higher limits. Economic nexus rules represent a fundamental shift toward destination-based taxation for e-commerce, aiming to realign sales tax remittance responsibilities with consumer locations. Building on these concerns, prior work has examined the implications of e-commerce taxation for business activity (Beem & Bruce, 2021), tax collection and incidence (Bruce et al., 2009; Fox et al., 2022), compliance behavior, and tax base erosion (Bruce & Fox, 2000). Still, the selection of threshold levels appears somewhat arbitrary, and additional research is needed to understand their economic and behavioral consequences.

In this paper, we estimate the impact of economic nexus threshold exposure on state-level business activity. Rather than treating all states with the same nominal threshold as equivalent, we construct a novel treatment measure—*threshold exposure*—defined as the economic nexus threshold amount scaled by state population (per 1,000 inhabitants). This measure captures the degree to which a given threshold is binding relative to the size of the local market: a $100,000 threshold is far more restrictive in a populous state like Florida than in a small state like South Dakota, where the same dollar amount exempts a much larger share of potential online sellers from tax collection obligations. Exploiting cross-state variation in threshold exposure following *Wayfair*, we test whether states where the threshold is less binding—leaving more online sellers exempt—experience distinct economic consequences for business activity compared to states where the threshold is more restrictive.

We employ a two-way fixed effects (TWFE) model as our baseline empirical strategy, estimating the impact of threshold exposure on various business outcomes such as number of firms and establishments, level of employment, and applications for new businesses. The model includes a rich set of time-varying controls identified as relevant in the sales tax literature—including demographic characteristics, income levels, internet penetration, labor market conditions, and tax rates—along with state and year fixed effects to absorb unobservable confounders. To assess the robustness of this baseline approach, we implement several alternative specifications: we rescale the treatment variable using goods-producing GDP rather than population, to address potential confounding from demographic movements; we gradually add controls in groups to evaluate coefficient stability and assess the potential for omitted variable bias; and we freeze treatment intensity at its first observed value and harmonize adoption timing to 2019, assessing whether staggered rollout or within-state variation in exposure drives the baseline results.

This work contributes to the literature on digital taxation and sales tax enforcement. At the broader level, *Wayfair* provides a compelling natural experiment to study how increased sales tax enforcement shapes business activity. By closing the

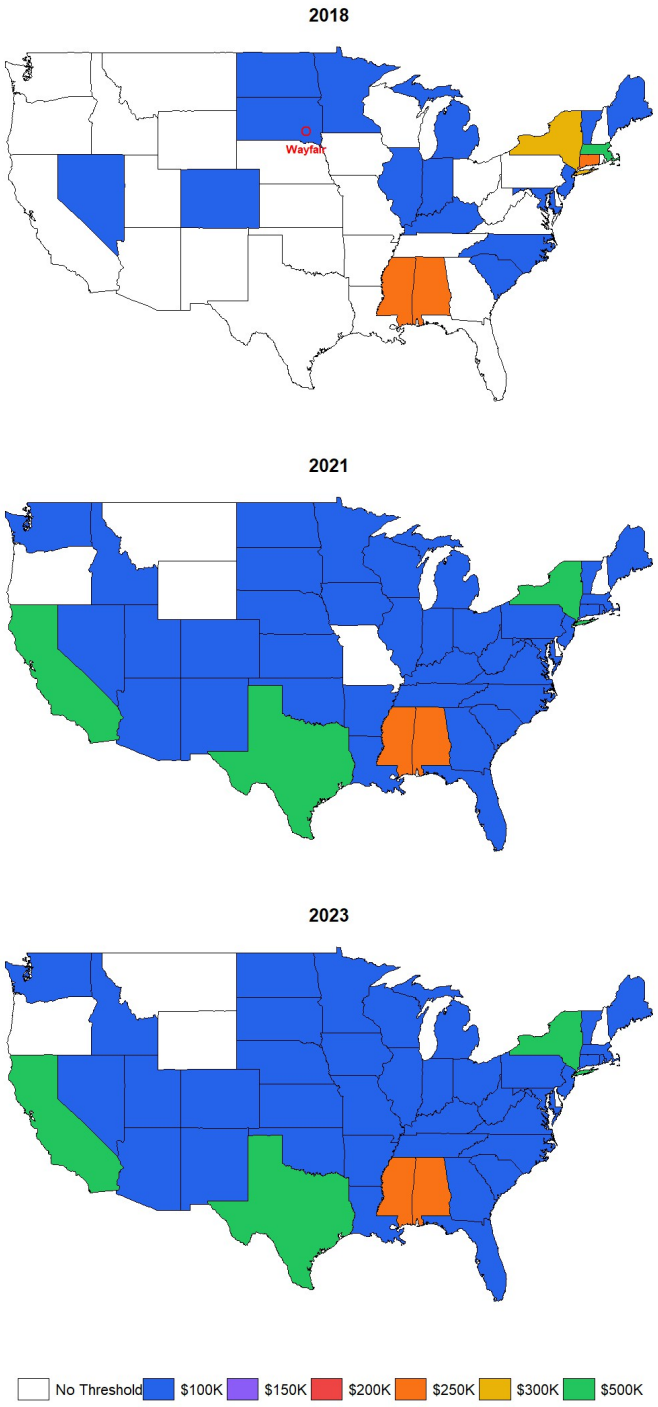


Figure 4.1: Adoption of Economic Nexus Thresholds Over Time

Notes: Each panel depicts the cumulative adoption of economic nexus policies across U.S. states at different points in time following the Supreme Court's ruling in *South Dakota v. Wayfair* (2018). Shading indicates whether a state had enacted an economic nexus threshold by the corresponding period, illustrating the rapid and near-universal diffusion of destination-based sales tax obligations for remote sellers in the years after *Wayfair*. *Source:* Sales Tax Institute.

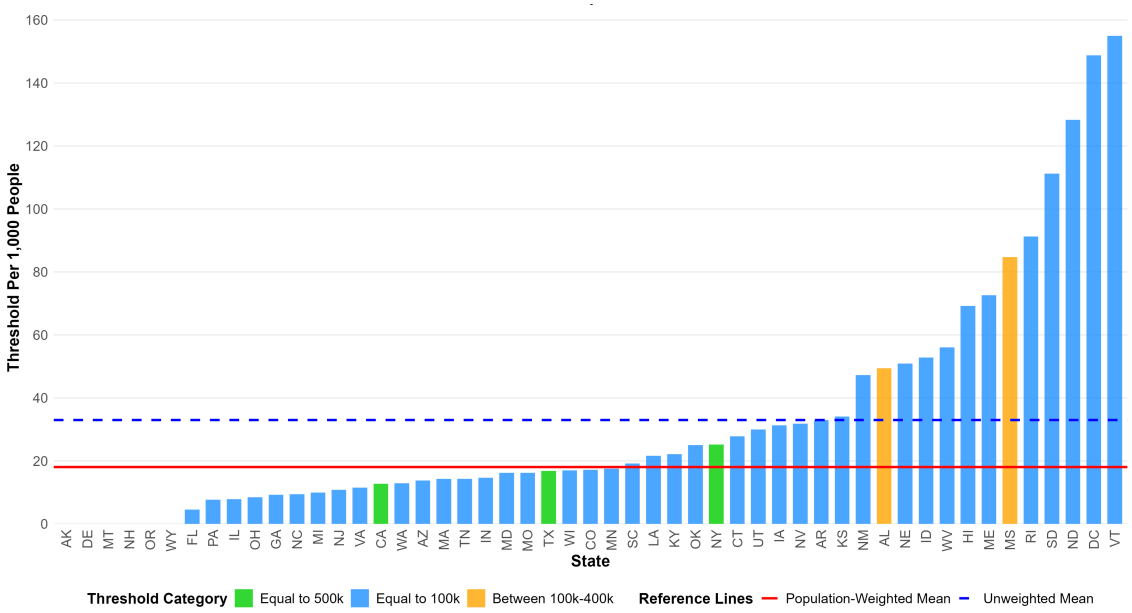


Figure 4.2: Economic Nexus Threshold Exposure Across U.S. States (2023)

Notes: This figure displays *threshold exposure* for each U.S. state as of 2023, defined as the nominal economic nexus threshold scaled by state population (per 1,000 inhabitants). A higher value indicates a less binding threshold relative to local market size, meaning a larger share of remote online sellers remain exempt from sales tax collection obligations. The same nominal threshold — most commonly $100,000 in sales and/or 200 transactions — is substantially less restrictive in small-population states than in large ones, motivating the use of threshold exposure as the key treatment variable rather than the nominal threshold level alone.

long-standing tax advantage enjoyed by remote sellers, *Wayfair* represents a shift toward greater tax equity between online and brick-and-mortar retailers—a correction that prior work suggests should benefit local firm activity (Agrawal & Shybalkina, 2024; Beem & Bruce, 2021). At a more specific level, this paper is the first to examine how variation in economic nexus threshold exposure influences these outcomes. While Beem and Bruce (2021) find that nexus laws increase firm entry and employment overall, they do not use post-*Wayfair* data and thus do not assess heterogeneity across threshold exposure levels. Agrawal and Shybalkina (2024) emphasize the importance of remittance rules for enforcement and redistributive tax revenue, but do not distinguish among threshold designs or their binding nature relative to market size. Fox et al. (2014) and Fox et al. (2022) argue that destination-based systems can improve equity and efficiency if appropriately structured. To the best of our knowledge, no research has rigorously examined whether differences in economic nexus threshold exposure have measurable impacts on state-level economic dynamics such as firm entry and employment.

We find small but statistically significant negative effects of higher threshold exposure on the number of firms, establishments, and employment. A one-unit increase in threshold exposure is associated with approximately a 0.01% decline in both firms and establishments and a 0.02% decline in employment, with effects concentrated among very small firms (fewer than 20 employees), which bear the brunt of the competitive disadvantage created by tax-exempt online competitors. Business applications show a negative but statistically insignificant association with threshold exposure ($p > 0.05$). These results suggest that states where the nexus threshold is less binding relative to market size leave a larger share of online sellers exempt from tax collection obligations, widening tax-induced price differentials that disadvantage physical retailers. Contrary to the common argument that higher thresholds protect small businesses, the findings suggest the opposite: it is precisely small brick-and-mortar firms that suffer most when the threshold is set too high relative to market size, shielding online competitors from tax parity. Threshold exposure—not merely

threshold level—is therefore the relevant policy parameter for ensuring competitive equity between remote and local commerce.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature and legal background. Section 3 presents the conceptual framework. Section 4 describes the data and empirical strategy. Section 5 presents the main results and heterogeneity analysis. Section 6 concludes with policy implications and directions for future research.

4.2 Legal Background & Literature

The expansion of internet access and digital infrastructure over the past three decades has reshaped commercial activity, with e-commerce emerging as a major channel for both business-to-consumer (B2C) and business-to-business (B2B) transactions. By 2018, online sales accounted for nearly 10% of total consumer sales and over 90% of B2B sales in the U.S. (Agrawal & Shybalkina, 2024).

This digital transformation has sparked growing concern among scholars and policymakers over the sustainability of traditional sales tax regimes and created competitive disparities between online and brick-and-mortar retailers. Early work by Bruce and Fox (2000) identified that e-commerce could function as a de facto tax haven, allowing consumers to bypass local tax obligations due to the absence of physical presence requirements. Remote vendors operating without physical nexus in a jurisdiction enjoyed significant sales tax exemptions, while traditional establishments faced corresponding price disadvantages that could reach approximately 10% in certain areas (Beem & Bruce, 2021). This asymmetric tax structure produced incentives for retailers to minimize their physical footprint to avoid tax obligations, while local governments often responded by lowering tax rates in an attempt to retain their tax base, effectively treating the internet as competitive pressure on fiscal policy (Agrawal, 2021).

The cumulative effect of these dynamics extended beyond competitive imbalances to generate measurable fiscal consequences. Beem and Bruce (2021) demonstrate that this behavior coincided with a substantial narrowing of the sales tax base, while fragmented enforcement mechanisms contributed to declines in both firm activity and state tax revenues. Conservative projections estimate that these untaxed electronic commerce transactions resulted in cumulative revenue shortfalls of $52 billion between 2007 and 2012 for state and local governments, highlighting the significant fiscal impact of the digital retail transformation.

The legal framework surrounding sales taxation of remote commerce in the United States has evolved substantially over the past three decades, with significant implications for state fiscal capacity. For years, the Supreme Court's decision in *Quill Corp. v. North Dakota* (1992) prohibited states from compelling out-of-state vendors to remit sales taxes unless they had a physical presence within the state. This precedent limited states' ability to enforce tax compliance on growing volumes of online sales, effectively enabling widespread avoidance of use taxes by consumers. The landscape shifted with the landmark decision in *South Dakota v. Wayfair, Inc.* (2018), in which the Court overturned *Quill* and authorized states to impose tax obligations based on economic presence. This ruling validated the use of economic nexus thresholds—typically defined by sales revenue or transaction counts—granting states the authority to require remote sellers to collect and remit taxes even without physical operations. The move toward destination-based sourcing and economic nexus represents an institutional reform that enhances the enforceability of sales taxes in a digital economy while redistributing tax capacity across jurisdictions (Agrawal & Shybalkina, 2024; Fox et al., 2014, 2022).

The relationship between taxation and entrepreneurial activity has been a subject of extensive research, though empirical findings remain mixed. Bruce and Schuetze (2004) and Bruce et al. (2020) reveal that while theoretical models suggest taxation should discourage entrepreneurship by reducing after-tax returns, empirical estimates of taxation's effect on entrepreneurship levels are generally inconclusive, with studies

finding positive, negative, or no significant relationships depending on the tax type, measurement approach, and economic context examined. This inconsistency persists across different tax instruments and entrepreneurship measures, highlighting the complex ways in which tax policies simultaneously influence both the costs and benefits of business formation. Beyond entry decisions, taxation also affects the quality and survival prospects of new ventures, Asoni and Sanandaji (2014) demonstrate that higher tax rates can improve the average quality of entrepreneurship by deterring marginal entrepreneurs while encouraging only those with stronger business prospects to enter the market, though this selection effect comes at the cost of reduced overall entrepreneurial activity.

Analyzing e-commerce in a pre-*Wayfair* environment, Bruce et al. (2015) examine how state tax policies influence online retailers' nexus establishment decisions using data from the top 500 e-commerce firms from 2006-2012. Their findings indicate that while firms prioritize market size over tax considerations when establishing nexus, higher sales tax rates and broader tax bases reduce the likelihood of nexus establishment, and states with higher e-tailer nexus experience greater retail employment benefits only when sales tax rates remain low.

Empirical evidence post-*Wayfair* demonstrates the varied effects of these legal changes on tax collection patterns and firms' decisions to allocate. Bruce et al. (2023) examine the effects of the *Wayfair* decision using county-level data from Tennessee and surrounding states alongside state-level business application data. Their analysis reveals that while *Wayfair* improved states' ability to enforce destination-based sales tax collection and reduced tax competition with previously untaxed internet commerce, traditional cross-border shopping competition has re-surged. Rural counties experienced greater sales tax collection growth while border counties saw reduced growth, and states with higher sales tax rates attracted more business applications post-*Wayfair* due to the elimination of nexus-establishment disincentives. In the same line, Agrawal and Shybalkina (2024) find that the adoption of economic nexus and marketplace facilitator laws increased local sales tax revenues

by 5.4 and 5.1 percent, respectively, with benefits accruing disproportionately to smaller and rural jurisdictions. This legal change corrected a long-standing asymmetry that had favored urban centers with retail agglomerations and contributed to tax base erosion. These findings collectively suggest that although *Wayfair* represented a step toward comprehensive destination-based taxation, the design of remittance rules is not neutral but shapes both the magnitude and spatial distribution of tax revenues, while interstate sales tax competition persists through cross-border shopping.

From the consumer perspective, research shows that consumer demand for e-commerce is highly responsive to tax policy. Ellison and Ellison (2009) provide early evidence that higher offline sales tax rates drive consumers toward online purchases, even when those online purchases are not directly taxed. Using detailed transaction-level data, they demonstrate that states with higher traditional retail tax rates tend to generate more online orders, suggesting that consumers actively seek tax avoidance opportunities. This behavior creates competitive distortions and highlights the role of offline taxes in shaping digital purchasing decisions. Extending this insight, Einav et al. (2014) estimate that a one-percentage-point increase in a state's sales tax results in a 2% increase in out-of-state online purchases and a decline of up to 4% in in-state purchases. This high elasticity underscores the importance of harmonizing tax treatment across jurisdictions to minimize distortions and promote fairness between online and offline commerce.

Contrary to concerns about overburdening small businesses, the enforcement of economic nexus standards did not deter economic activity. On the contrary, some jurisdictions experienced increases in business formation and employment as the tax system became more neutral (Agrawal & Shybalkina, 2024; Beem & Bruce, 2021). Yet despite this convergence in adoption and enforcement, states differ meaningfully in how binding their thresholds are relative to local market size. Among states sharing the same nominal threshold, the effective stringency of the policy varies considerably: a $100,000 threshold represents a far more restrictive constraint in a large, populous

state than in a small one, where the same dollar amount exempts a disproportionately large share of online sellers from tax collection obligations. This variation in *threshold exposure*—the threshold amount scaled by state population—raises questions about whether the current threshold design achieves competitive neutrality between online and offline retailers, and whether its economic consequences differ systematically across states with different market sizes.

4.3 Conceptual Framework

Threshold levels may shape economic behavior by altering the incentives that firms and consumers face. From the firm's perspective, economic nexus thresholds generate opposing forces on business formation and survival. For prospective entrepreneurs, the compliance burden of tracking sales across multiple jurisdictions and remitting taxes to various state authorities raises the costs of operating online, potentially deterring e-commerce entry. At the same time, the post-*Wayfair* environment may encourage traditional retail entry by leveling a competitive playing field that previously tilted toward remote sellers who could offer tax-free prices.

Among established firms, particularly those located in no-sales-tax states that previously competed on tax-exempt pricing, economic nexus requirements compress margins and may force exit from markets where demand cannot absorb the new tax burden. The severity of this pressure depends heavily on where a firm's customers are located: firms serving consumers in high-tax states face sharper demand responses than those serving low-tax markets. Firms that survive the transition, however, face weakened incentives to cluster operations in tax-advantaged states, since physical presence no longer determines tax obligations. This may encourage geographic expansion of establishments closer to customer bases, offsetting some of the negative effects on firm counts and employment. As these forces work in opposing directions (see Figure 4.3), the net effect on aggregate business activity—firm counts, establishment numbers, and employment—is theoretically ambiguous.

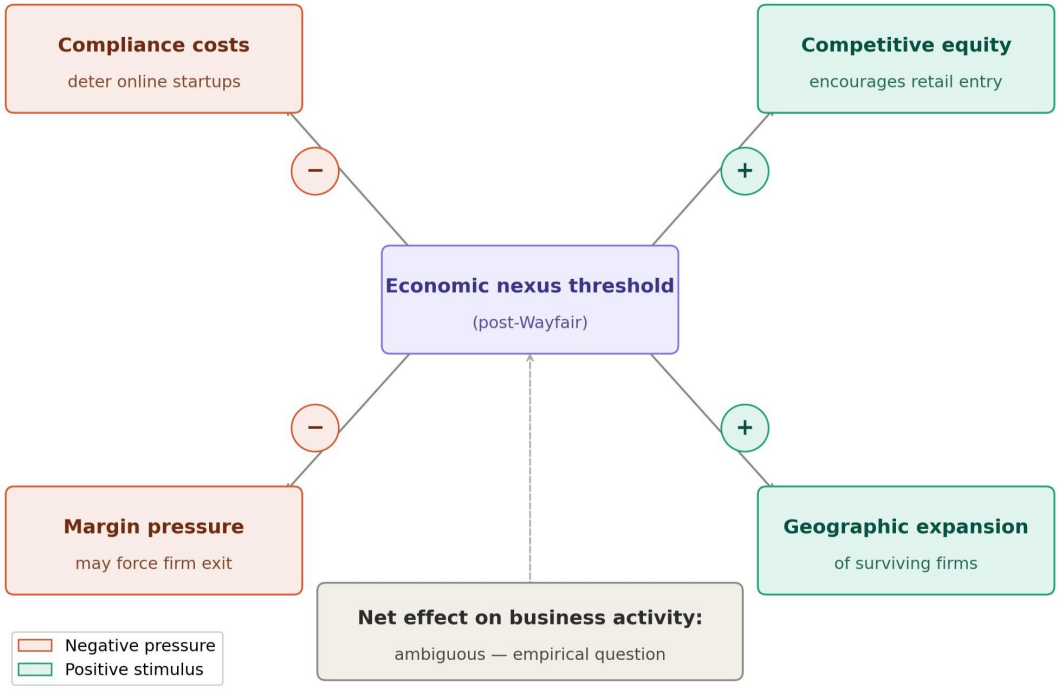


Figure 4.3: Competing Effects of Economic Nexus Thresholds on Business Activity

Notes: This figure illustrates the opposing forces that economic nexus thresholds generate on business activity. Higher threshold exposure shields more remote sellers from tax collection obligations, deterring traditional retail entry and compressing margins for incumbent firms competing on tax-exempt pricing, while simultaneously reducing compliance costs for online entrants and weakening incentives to cluster operations in tax-advantaged states — leaving the net effect on firm counts, establishments, and employment theoretically ambiguous.

From the consumer's perspective, tax equalization reduces the price advantage that previously made online shopping attractive to buyers in high-tax states, potentially redirecting purchasing toward local storefront retailers (Agrawal & Shybalkina, 2024; Beem & Bruce, 2021). The reach of marketplace facilitator laws reinforces this shift by closing the loophole through which consumers could avoid taxes by buying from small vendors on major platforms, making tax-free online purchases increasingly difficult regardless of seller size.

4.4 Empirical Framework

4.4.1 Data Sources

This study draws on a state-level panel covering the period from 2013 to 2023, providing five years of data both before and after the *Wayfair* decision.

Business Outcomes. The primary source for business formation and entrepreneurial dynamics is the *Business Formation Statistics* (BFS) from the U.S. Census Bureau, which provides monthly, state-level data on new business applications and formations. To complement this analysis and examine impacts across different firm types, sizes, and ages beyond startups, we incorporate data from the *Business Dynamics Statistics* (BDS), also from the U.S. Census Bureau, which tracks the number of firms, establishments, and employment levels across age and size categories.

Treatment Variable. We obtain information on economic nexus threshold levels and adoption dates from the *Sales Tax Institute*, a third-party source that tracks state-level sales tax legislation. Following *Wayfair*, states generally converged around a $100,000 sales threshold, though California, New York, and Texas maintained higher $500,000 thresholds, and a small number of states adopted intermediate values. Our treatment variable—*threshold exposure*—is constructed as the economic nexus threshold amount (in dollars) divided by state population (per 1,000 inhabitants),

capturing the degree to which the threshold is binding relative to the size of the local market rather than treating all states with the same nominal threshold as equivalent.

Control Variables. The selection of control variables follows Beem and Bruce (2021), who emphasize several important forces behind changes in the sales tax base that should be accounted for when isolating the effect of nexus standards. In particular, we proxy for the rise of remote sales and the shift in consumption toward services, and we supplement these with additional controls for broader economic conditions.

To capture the rise of internet retail and consumers' ability to access online marketplaces, we include a measure of household high-speed internet connections from the *Federal Communications Commission* (FCC). To account for the shift away from goods-producing industries toward services, we control for the service-providing share of the economy using data from the *Bureau of Economic Analysis* (BEA); failure to include this would bias estimates depending on the relative growth rates of service versus goods-producing firms. To address the role of extractive industries—which allow states to export part of their sales tax burden to out-of-state taxpayers—we control for oil and coal production using data from the *State Energy Data System* (SEDS) of the Energy Information Administration.

Beyond these tax-base controls, we include a rich set of additional time-varying state characteristics that may independently influence business dynamics. From the *American Community Survey* (ACS), we obtain population, median household income, the income share of the top 5% of residents*—to account for the income distribution's impact on the sales tax base and business activity—college degree attainment rates, and the share of non-white residents. The unemployment rate from the *Bureau of Labor Statistics* (BLS) serves as a business cycle control. Per-capita government spending from the *Bureau of Economic Analysis* (BEA) and maximum corporate and personal income tax rates from the *Tax Foundation* account

*The estimated income share of the top 5% may be understated because ACS income variables are subject to top-coding, which compresses variation in the upper tail of the income distribution. See https://forum.ipums.org/t/top-coding-incwage-in-5-year-acs-samples/5328.

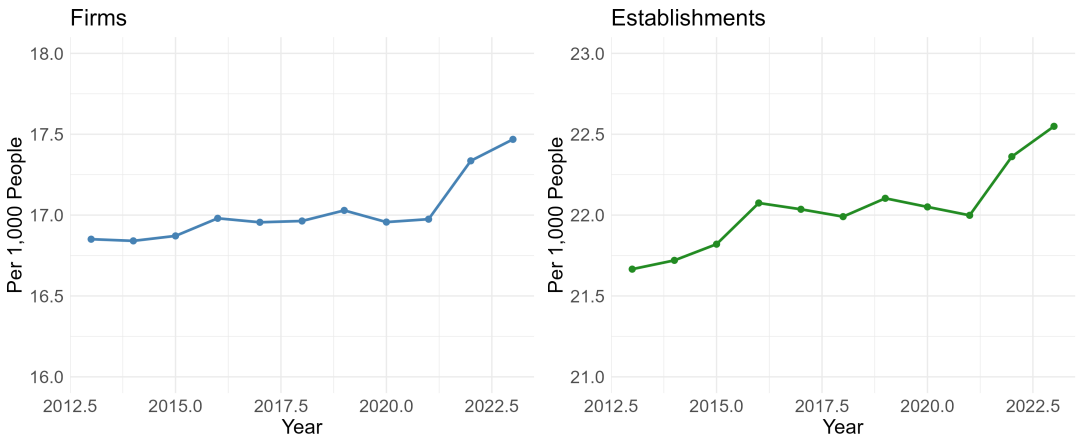
for other dimensions of state fiscal policy that may independently affect business formation and location decisions. Finally, property crime rates from the *Uniform Crime Reports* (UCR) are included given that perceived safety is a relevant factor in business formation and expansion decisions.

4.4.2 Descriptive Evidence

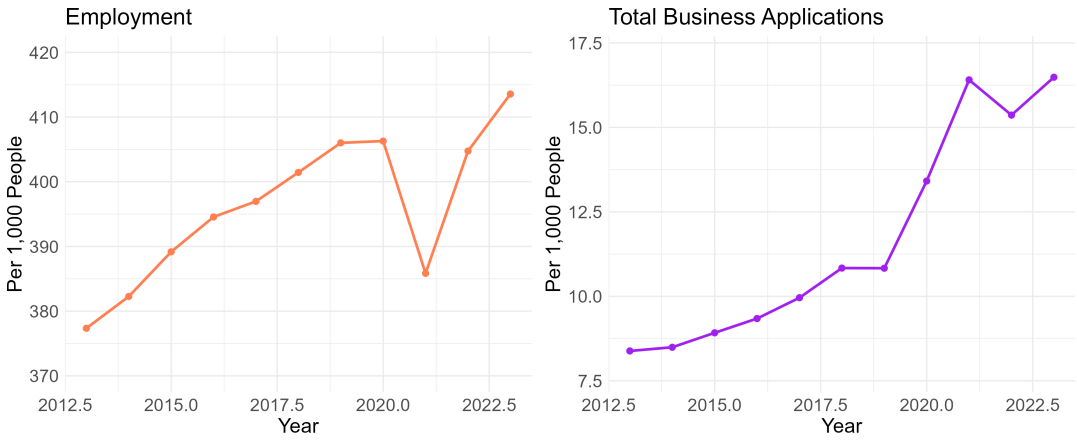
Figure 4.1 documents the rapid and widespread adoption of economic nexus thresholds following the *Wayfair* decision. By 2020, the vast majority of U.S. states had converged toward a $100,000 threshold, with only a handful of large states—California, New York, and Texas—maintaining the higher $500,000 standard. Figure 4.2 illustrates the resulting variation in threshold exposure across states in 2023: even among states sharing the same nominal $100,000 threshold, the effective tightness of the policy varies considerably once scaled by population, with smaller states exhibiting substantially higher exposure values than larger ones. This cross-state variation in threshold exposure—rather than simple threshold adoption—is the key source of identification exploited in the empirical analysis.

Turning to aggregate trends, Figures 4.4a and 4.4b present the national evolution of four outcome variables over the sample period. Business applications grew steadily from roughly 8.5 per 1,000 people in 2013 to over 16 per 1,000 people by 2023, with a particularly sharp acceleration after 2019. The number of firms, establishments, and employment levels per 1,000 people also trended upward over the sample period, though more gradually, with firms rising from around 16.9 to 17.5, establishments from 21.7 to 22.6, and employment from approximately 377 to 413 per 1,000 people.

These descriptive patterns are inherently limited in their ability to identify causal effects: the post-*Wayfair* period coincides closely with the COVID-19 pandemic, which independently disrupted business formation and consumer behavior in ways that are difficult to disentangle from the policy changes of interest. This limitation motivates the econometric approach developed in the following section, which exploits



(a) Firms and Establishments (per 1,000 People)



(b) Employment and Total Business Applications (per 1,000 People)

Figure 4.4: Overall Business Activity Trends in the United States, 2013–2023

cross-state variation in threshold exposure to separately identify the effects of nexus threshold design from these confounding aggregate shocks.

4.4.3 Econometric Approach

Our baseline empirical strategy employs a two-way fixed effects (TWFE) model to estimate the effects of economic nexus threshold exposure on business outcomes. The main specification is:

$$\text{Outcome}_{st} = \alpha_s + \tau_t + \beta \cdot \text{Threshold_Exp}_{st} + \mathbf{X}'_{st}\delta + \epsilon_{st} \tag{4.1}$$

where Outcome_{st} represents various outcome of interest for state s in year t, including total business applications, number of firms, establishments and employment. $\text{Threshold_Exp}_{st}$ is the key treatment variable, defined as the economic nexus threshold amount scaled by state population (per 1,000 inhabitants), capturing the degree to which the threshold is binding relative to local market size. α_s and τ_t are state and year fixed effects, respectively, absorbing time-invariant state characteristics and common aggregate shocks. $\mathbf{X}_{st}$ is a vector of time-varying state controls including population, median income, top 5% income share, college degree attainment, share of non-white residents, unemployment rate, service sector share of the economy, per-capita government spending, property crime rates, high-speed internet penetration, maximum income and corporate tax rates, and oil and coal production. Standard errors are clustered at the state level.

Identification and Threats. The TWFE estimator relies on two key assumptions. First, because threshold exposure is a continuous treatment, identification of dose-response comparisons across dose levels requires the strong parallel trends condition of Callaway et al. (2024), which requires that states with different levels of threshold exposure would have followed similar outcome trajectories absent the policy variation *and* that the level of dose is not systematically related to pre-existing differences in outcome trends. Second, the exogeneity assumption requires

that threshold levels are not set in anticipation of future business outcomes. While state and year fixed effects absorb a broad set of confounders, several threats remain. The most significant is the presence of concurrent policy changes—most notably the COVID-19 pandemic, which coincided closely with the post-*Wayfair* adoption period and independently affected business formation and tax revenues across all states. Additionally, time-varying state characteristics correlated with both threshold choice and business outcomes could bias estimates if not adequately controlled for. The rich set of controls described above is designed to mitigate these concerns, though the counterfactual—what would have happened to business activity in each state absent the specific threshold adopted—remains inherently unobservable.

We address these concerns by exploiting cross-state variation in threshold *exposure* rather than the aggregate pre-post contrast, providing identification from within-period differences across states rather than relying on a before-after comparison that conflates the policy change with contemporaneous aggregate shocks such as the pandemic. Regarding exogeneity, threshold levels were unlikely to be chosen in anticipation of specific business outcomes. Rather than conducting rigorous cost-benefit analyses, most states engaged in a discretionary "follow-the-leader" process after the *Wayfair* ruling, adopting the $100,000 standard pioneered by South Dakota with little evidence of systematic tailoring to state-specific economic conditions. This process, driven by legal precedent and legislative convenience rather than anticipated economic returns, supports treating threshold exposure as plausibly exogenous. To the extent that the strong parallel trends condition of Callaway et al. (2024) holds—consistent with the discretionary nature of threshold adoption—our TWFE estimates recover a meaningful weighted average of dose-specific treatment effects.

4.5 Results

We present the baseline TWFE estimates of the effect of threshold exposure in Table 4.1. Across columns, higher threshold exposure is associated with small but

statistically significant negative effects on the number of firms, establishments, and employment, while business applications show no significant response.

The point estimates on firms and establishments are identical at -0.0001 and statistically significant at the 10% level, while the employment effect is larger in magnitude at -0.0002 and significant at the 5% level. To interpret the magnitudes, consider first a comparison between South Dakota and Iowa in 2023. South Dakota, with its small population and $100,000 threshold, exhibits a threshold exposure of approximately $110 per 1,000 inhabitants, while Iowa's larger population implies an exposure of roughly $35 per 1,000 inhabitants—a difference of 75 units. Multiplying this difference by the employment coefficient yields an implied employment gap of $75 \times (-0.0002) = -0.015$, or approximately **1.5% lower employment** in South Dakota relative to Iowa associated with the difference in threshold exposure.

More generally, a movement from the 75th percentile (looser threshold) to the 25th percentile (tighter threshold) of threshold exposure in 2023—from 40.7 to 10.4 per 1,000 inhabitants, a tightening of 30.3 units—is associated with a $30.3 \times (-0.0001) = 0.003$, or roughly **0.3% more firms and establishments**, and a $30.3 \times (-0.0002) = 0.006$, or approximately **0.6% higher employment**. While modest in absolute terms, these effects are economically meaningful given that they likely operate through the competitive channel of tax-exempt online sellers[†].

4.5.1 Heterogeneity by Size and Age

Table 4.2 breaks down the main results by firm size. The negative effects are driven by very small firms (fewer than 20 employees), while larger firms are largely unaffected.

Panel A shows significant negative effects on firms, establishments, and employment for very small firms, all at the 5% level. These are the businesses most exposed to price competition from online sellers that remain below the tax collection threshold.

[†]This work does not quantify welfare losses from interstate misallocation. Instead, it examines how variation in economic nexus threshold exposure affects within-state business activity through changes in tax parity between online and offline retailers. The estimates should therefore be interpreted as reduced-form state-level effects, not as evidence on interstate reallocation.

Table 4.1: Impact of Economic Nexus Threshold Exposure

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)	Log(BA) (4)
Threshold Exposure	−0.0001*	−0.0001*	−0.0002**	−0.0007
	(0.0001)	(0.0001)	(0.0001)	(0.0005)
State & Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	557	557	557	557
Within R^2	0.728	0.762	0.749	0.225

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at state level in parentheses. Threshold Exposure is the economic nexus threshold (in dollars) normalized by state population (per 1,000 inhabitants). Controls include: population, median income, top 5% income, college degree %, non-white %, unemployment rate, service sector share, government spending per capita, crime rates, high-speed internet, income tax rate, corporate tax rate, coal production, and oil production. *p<0.1; **p<0.05; ***p<0.01

Table 4.2: Impact of Threshold Exposure by Firm Size

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)
Panel A: Very Small Firms (<20 employees)			
Threshold Exposure	−0.0002**	−0.0002**	−0.0002**
	(0.0001)	(0.0001)	(0.0001)
Panel B: Small Firms (20–499 employees)			
Threshold Exposure	−0.0001	−0.0001	−0.0002
	(0.0001)	(0.00005)	(0.0001)
Panel C: Medium Firms (500–2,499 employees)			
Threshold Exposure	0.0003***	0.0001	−0.0003*
	(0.0001)	(0.0002)	(0.0002)
Panel D: Large Firms (2,500+ employees)			
Threshold Exposure	0.0001	0.0002*	−0.0001
	(0.0001)	(0.0001)	(0.0002)
State & Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at state level in parentheses. All specifications include the full set of controls as in Table 4.1. *p<0.1; **p<0.05; ***p<0.01

Larger firms (Panels B–D) show no consistent negative effects, and medium firms even show a positive effect on firm counts. Bigger firms can more easily absorb price differentials, or operate across multiple channels. This result cuts against the argument that higher thresholds protect small businesses—in practice, it seems that small local retailers are paying the price when the threshold is set too high.

Table 4.3 repeats the analysis by firm age. The estimates are negative across all groups but statistically insignificant throughout.

The largest point estimates are for startups and young firms (Panels A and B), with employment coefficients of -0.0007 and -0.0005, but neither is significant. The age dimension does not provide clean identification, and the results here are best read as suggestive rather than conclusive.

4.5.2 Heterogeneity by Industry Type

Next, we examine heterogeneity by industry. One might expect that sectors with greater online exposure—such as retail trade or wholesale trade—would experience stronger effects, as firms in these industries face more direct competition from tax-exempt online sellers.

The results in Table 4.4 reveal meaningful heterogeneity across industries. The most consistently impacted sectors are Retail Trade and Transportation and Warehousing. This pattern is broadly consistent with Beem and Bruce (2021), who find that retail trade and related sectors are most responsive to sales tax nexus non-neutralities. Retail and logistics firms operate in direct competition with remote online sellers, so higher thresholds—which allow more online sellers to avoid collecting sales tax—translate most immediately into a price disadvantage for local business. Sectors such as construction, administrative services, and healthcare show small and statistically insignificant effects, suggesting that the competitive distortion introduced by economic nexus thresholds is concentrated in the traded-goods and distribution segments of the economy.

Table 4.3: Impact of Threshold Exposure by Firm Age

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)
Panel A: Startups (0 years)			
Threshold Exposure	−0.0003	−0.0003	−0.0007
	(0.0003)	(0.0003)	(0.0004)
Panel B: Young Firms (1–5 years)			
Threshold Exposure	−0.0003	−0.0003	−0.0005
	(0.0002)	(0.0002)	(0.0004)
Panel C: Established Firms (6–10 years)			
Threshold Exposure	−0.00002	−0.00004	−0.0002
	(0.0002)	(0.0001)	(0.0002)
Panel D: Mature Firms (11+ years)			
Threshold Exposure	−0.00004	−0.00002	−0.0001
	(0.00005)	(0.00005)	(0.0001)
State & Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at state level in parentheses. All specifications include the full set of controls as in Table 4.1. *p<0.1; **p<0.05; ***p<0.01

Table 4.4: Estimation Results by Industrial Sector

	ln(Firms) (1)	ln(Establishments) (2)	ln(Employment) (3)
Mining, Quarrying, and Oil and Gas Extraction	0.0006	0.0004	−0.0002
	(0.0004)	(0.0003)	(0.0011)
Utilities	−0.0004*	−0.0005*	−0.0003
	(0.0002)	(0.0003)	(0.0005)
Construction	−0.0001	−0.0001	−0.0001
	(0.0001)	(0.0001)	(0.0002)
Manufacturing	−0.0000	−0.0000	−0.0001
	(0.0001)	(0.0001)	(0.0001)
Wholesale Trade	−0.0001	−0.0001	−0.0003
	(0.0001)	(0.0001)	(0.0002)
Retail Trade	−0.0002***	−0.0001	−0.0002**
	(0.0001)	(0.0001)	(0.0001)
Transportation and Warehousing	−0.0004**	−0.0004**	−0.0007*
	(0.0002)	(0.0002)	(0.0003)
Information	−0.0005	−0.0002	−0.0002
	(0.0003)	(0.0003)	(0.0002)
Real Estate and Rental and Leasing	−0.0002	−0.0002	−0.0002
	(0.0001)	(0.0001)	(0.0002)
Professional, Scientific, and Technical Services	−0.0002	−0.0003	−0.0007
	(0.0002)	(0.0002)	(0.0004)
Management of Companies and Enterprises	−0.0001	−0.0002	0.0004
	(0.0002)	(0.0003)	(0.0005)
Administrative and Support and Waste Management	0.0001	0.0001	0.0000
	(0.0001)	(0.0001)	(0.0002)
Educational Services	−0.0001	−0.0000	−0.0002
	(0.0001)	(0.0001)	(0.0002)
Health Care and Social Assistance	−0.0002	−0.0002	−0.0001
	(0.0001)	(0.0001)	(0.0001)
Arts, Entertainment, and Recreation	−0.0003*	−0.0003**	−0.0001
	(0.0001)	(0.0001)	(0.0003)
Accommodation and Food Services	−0.0001	−0.0001	−0.0003**
	(0.0001)	(0.0001)	(0.0001)
Other Services	−0.0000	−0.0000	−0.0001
	(0.0001)	(0.0001)	(0.0001)
Agriculture, Forestry, Fishing and Hunting	−0.0002	−0.0001	0.0000
	(0.0002)	(0.0002)	(0.0004)
Finance and Insurance	−0.0000	0.0002*	0.0001
	(0.0001)	(0.0001)	(0.0002)
State & Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Each row presents the coefficient on Threshold Exposure (economic nexus threshold normalized by state population per 1,000 inhabitants) from a separate panel regression with state and year fixed effects and the full control set from Table 4.1. Standard errors clustered at the state level in parentheses. *p<0.1; **p<0.05; ***p<0.01

4.6 Alternative Specifications

4.6.1 Fixed Population Denominator

A concern with the baseline treatment measure is that, because it is constructed as the statutory threshold divided by state population, its post-*Wayfair* variation is driven largely by cross-state differences in market size. Since most states converged to a common $100,000 threshold after *Wayfair*, the identifying variation in threshold exposure comes primarily from differences in population, so that in a two-way fixed effects setting the treatment is closely related to a post-*Wayfair* interaction with inverse population. To address this concern, we construct an alternative treatment variable by dividing each state's economic nexus threshold by its 2017 population and holding that denominator fixed throughout the panel. Using 2017—the last full pre-*Wayfair* year—preserves the cross-state variation in preexisting market size while removing any influence of post-treatment population changes from the treatment measure.

Table 4.5 presents the estimates using this fixed-population denominator. The estimated coefficients remain negative across all outcomes. Relative to the baseline specification in Table 4.1, the establishment coefficient loses precision, while the coefficients for firms and employment remain statistically significant. The similarity of these estimates suggests that the baseline findings are not driven by contemporaneous population movements.

4.6.2 Market Size Alternative Measure

As a further robustness check, we replace population with a fixed measure of the size of each state's goods market based on pre-*Wayfair* GDP. Specifically, for each state we use the dollar value of GDP originated by goods-producing industries associated with private consumption—excluding investment and government spending—and average this measure over 2015–2017. We then express this denominator in *millions of*

dollars and divide the state's statutory economic nexus threshold (in dollars) by that fixed amount. The resulting treatment variable is therefore interpreted as threshold dollars per $1 million of pre-*Wayfair* goods-producing consumption GDP. Using the 2015–2017 average fixes the denominator prior to *Wayfair* and smooths short-run fluctuations in state economic activity.

Figure 4.5 displays the cross-state variation in this alternative treatment measure for 2023. Table 4.6 presents the corresponding regression results. The estimates are consistent with the baseline findings in Table 4.1, with the threshold coefficient remaining negative and statistically significant for firms, establishments, and employment, while business applications remain statistically insignificant. Overall, the results are robust to scaling the threshold by a fixed pre-*Wayfair* measure of goods-market size rather than by population.

4.6.3 Gradual Addition of Controls

To explore whether the estimated effect of economic nexus threshold exposure on measures of business activity exhibits stability with the incremental inclusion of controls, we turn to Table 4.7. For each outcome (i.e., firms, establishments, and employment), we present treatment effects for threshold exposure using only state and year fixed effects and three groups of controls added incrementally to the estimating equation. If the estimated effect of nexus on business activity is stable across different specifications, then omitted variable bias is unlikely to be a concern.

Table 4.7 shows that the coefficient on threshold exposure is stable across specifications. For firms and establishments it ranges from -0.0002 (no controls) to -0.0001 (full control set), while for employment it moves from -0.0003 to -0.0002, retaining significance throughout. The modest change in magnitude as controls are added incrementally suggests that omitted variable bias is unlikely to drive the baseline results.

Table 4.5: Alternative Treatment: Threshold Scaled by Fixed 2017 Population

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)	Log(BA) (4)
Threshold per 1,000 (Fixed Pop.)	−0.0001*	−0.0001	−0.0002**	−0.0007
	(0.0001)	(0.0001)	(0.0001)	(0.0005)
State & Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	557	557	557	557

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at the state level in parentheses. The alternative treatment variable is the economic nexus threshold (in dollars) normalized by state population measured in 2017 and held fixed across the panel, expressed per 1,000 inhabitants, so that the scaling factor is unaffected by post-treatment population changes. Controls include: population, median income, top 5% income, college degree %, non-white %, unemployment rate, service sector share, government spending per capita, crime rates, high-speed internet, income tax rate, corporate tax rate, coal production, and oil production. *p<0.1; **p<0.05; ***p<0.01

Table 4.6: Alternative Treatment: Threshold Scaled by Goods-Producing Consumption GDP

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)	Log(BA) (4)
Threshold ($) / GDP-Consumption ($ millions)	−0.0005*	−0.0005*	−0.0014***	−0.0004
	(0.0003)	(0.0003)	(0.0003)	(0.0022)
State & Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	557	557	557	557
Within R^2	0.726	0.761	0.758	0.211

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at the state level in parentheses. The alternative treatment variable is the economic nexus threshold (in dollars) normalized by a fixed measure of state market size: the average annual dollar value of GDP originated by goods-producing industries associated with private consumption (excluding investment and government spending), computed over 2015–2017 and expressed in millions of dollars. This pre-*Wayfair* denominator is held constant across the panel, so the treatment measure is interpreted as threshold dollars per $1 million of goods-producing consumption GDP and is unaffected by post-treatment economic or demographic changes. Controls include: population, median income, top 5% income, college degree %, non-white %, unemployment rate, service sector share, government spending per capita, crime rates, high-speed internet, income tax rate, corporate tax rate, coal production, and oil production. *p<0.1; **p<0.05; ***p<0.01

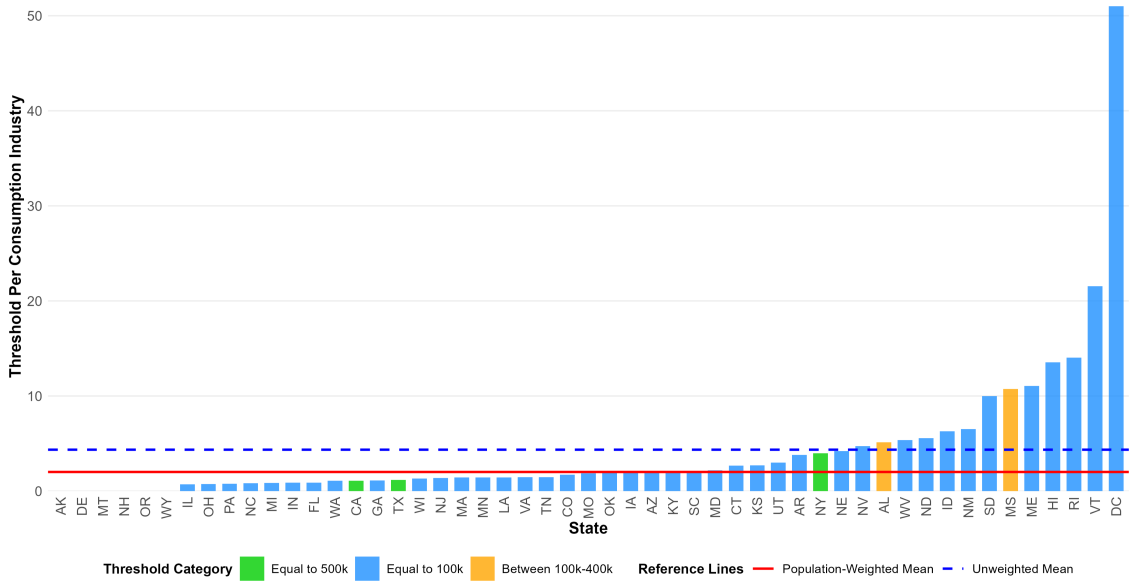


Figure 4.5: Threshold per $1 Million of Goods-Producing Consumption GDP Across States (2023)

Notes: The figure plots the alternative treatment variable for 2023, defined as each state's statutory economic nexus threshold (in dollars) divided by the state's average 2015–2017 goods-producing GDP associated with private consumption, expressed in millions of dollars. Higher values indicate that the threshold is less binding relative to the size of the state's pre-*Wayfair* goods market.

Table 4.7: Baseline Estimation Results with Gradual Inclusion of Controls

	ln(Firms)				ln(Establishments)				ln(Employment)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Threshold per 1,000 People	−0.0002**	−0.0002**	−0.0001	−0.0001**	−0.0002**	−0.0002**	−0.0001	−0.0001**	−0.0003***	−0.0003***	−0.0002**	−0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Within R^2	0.026	0.630	0.682	0.713	0.026	0.672	0.728	0.751	0.051	0.533	0.721	0.744
Observations	562	562	562	557	562	562	562	557	562	562	562	557
Controls (Group 1)		✓	✓	✓		✓	✓	✓		✓	✓	✓
Controls (Group 2)			✓	✓			✓	✓			✓	✓
Controls (Group 3)				✓				✓				✓
Fixed Effect (State)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed Effect (Year)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Group 1 controls: log(Population), Median Income, and broadband Internet connections. Group 2 controls: service-providing share of GDP, oil production, coal production, unemployment rate, and top 5% income share. Group 3 controls: government spending, top corporate tax rate, top personal income tax rate, property crime rate, and educational attainment. All regressions include state and year fixed effects. Sample period: 2013–2023. Standard errors clustered at the state level. *p<0.1; **p<0.05; ***p<0.01

4.6.4 Adoption Timing Check

Once South Dakota prevailed in *Wayfair* (2018), states faced immediate pressure to adopt economic nexus policies. However, not all states moved simultaneously—some enacted legislation in 2019, others waited until 2020 or later. This staggered rollout raises a well-known concern: two-way fixed effects estimators are not robust to treatment effect heterogeneity across adoption cohorts (Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021). If early and late adopters experienced systematically different effects, the baseline TWFE coefficient in Table 4.1 may be a weighted average that obscures or distorts the true treatment effect.

To assess whether this is a practical concern in this application, we implement a simple timing check. we artificially assign treatment onset to 2019 for all adopting states—the year immediately following the *Wayfair* decision—thereby eliminating variation in adoption timing. In addition, we freeze each state's treatment intensity at its first observed positive threshold value, holding it constant through 2023. This removes any influence of within-state changes in exposure over time. Pre-*Wayfair* years (2013–2018) retain a threshold of zero for all states. Table 4.8 reports the results under this alternative treatment setup.

The estimates are similar to the baseline results in Table 4.1 and suggest that staggered adoption timing and variation in treatment intensity over time are unlikely to be a meaningful source of bias in this setting—a finding consistent with the relatively compressed adoption window observed across most states following *Wayfair*.

4.7 Discussion

This work examines whether variation in economic nexus threshold exposure—measured as the threshold amount scaled by state population—affects business activity across U.S. states following the *Wayfair* decision. Using a state-level panel from 2013 to 2023 and a two-way fixed effects model with a rich set of controls, we

find that higher threshold exposure is associated with small but significant negative effects on the number of firms, establishments, and employment.

The heterogeneity results deepen the interpretation: the negative effects are concentrated among very small firms (fewer than 20 employees), which are the most exposed to price competition from online sellers that remain exempt from tax collection when thresholds are set high relative to market size. Larger firms show no consistent negative response. This pattern runs counter to the common policy argument that higher thresholds protect small businesses—the evidence here suggests the opposite. At the industry level, the effects are most pronounced in retail trade and transportation and warehousing, two sectors with significant exposure to online competition and logistics disruption, while most other industries show no significant response.

These findings have direct implications for threshold design. States should consider not just the nominal dollar value of their threshold but how binding that threshold is given their market size. A uniform $100,000 threshold does not create uniform competitive conditions across states: it is far more lenient in small states, where it exempts a larger share of online sellers from tax obligations and widens the price disadvantage faced by local retailers. Policymakers considering threshold harmonization or reform should account for this variation, as simply aligning nominal thresholds across states would leave meaningful differences in effective enforcement rules unaddressed. More broadly, the results contribute to the literature on sales tax enforcement and business activity (Agrawal & Shybalkina, 2024; Beem & Bruce, 2021; Bruce & Fox, 2000; Bruce et al., 2023; Fox et al., 2022) by showing that *how* economic nexus thresholds are designed—not just whether they exist—has measurable consequences for local economies.

References

Agrawal, D. R. (2021). The internet as a tax haven? *American Economic Journal: Economic Policy*, *13*(4), 1–35.

Agrawal, D. R., & Shybalkina, I. (2024). Remittance rules and the distribution of local tax revenue: Evidence after wayfair. *National Tax Journal*, *77*(4), 814–831.

Asoni, A., & Sanandaji, T. (2014). Taxation and the quality of entrepreneurship. *Journal of Economics*, *113*, 101–123.

Beem, R., & Bruce, D. (2021). Failure to launch: Measuring the impact of sales tax nexus standards on business activity. *Journal of Public Economics*, *201*, 104476.

Bruce, D., & Fox, W. F. (2000). E-commerce in the context of declining state sales tax bases. *National Tax Journal*, *53*(4), 1373–1388.

Bruce, D., Fox, W. F., & Luna, L. (2009). State and local government sales tax revenue losses from electronic commerce. *State Tax Notes*, *52*(7), 537–557.

Bruce, D., Fox, W. F., & Luna, L. (2015). E-tailer sales tax nexus and state tax policies. *National Tax Journal*, *68*(3S), 735–766. https://doi.org/10.17310/ntj.2015.3s.02

Bruce, D., Fox, W. F., & Shute, A. M. (2023). *Wayfair: A step towards the destination, but sales tax competition remains* (tech. rep.). National Bureau of Economic Research.

Bruce, D., Gurley-Calvez, T. J., Norwood, A., et al. (2020). Taxes and entrepreneurship: A literature review and research agenda. *Foundations and Trends® in Entrepreneurship*, *16*(5), 393–443.

Bruce, D., & Schuetze, H. J. (2004). Tax policy and entrepreneurship. *Swedish Economic Policy Review*, *11*(2), 233–265.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2024). *Difference-in-differences with a continuous treatment* (Working Paper No. 32117) (Also

available as arXiv:2107.02637). National Bureau of Economic Research. https://doi.org/10.3386/w32117

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Einav, L., Knoepfle, D., Levin, J., & Sundaresan, N. (2014). Sales taxes and internet commerce. *American Economic Review*, *104*(1), 1–26.

Ellison, G., & Ellison, S. F. (2009). Tax sensitivity and home state preferences in internet purchasing. *American Economic Journal: Economic Policy*, *1*(2), 53–71. https://doi.org/10.1257/pol.1.2.53

Fox, W. F., Hargaden, E. P., & Luna, L. (2022). Statutory incidence and sales tax compliance: Evidence from wayfair. *Journal of Public Economics*, *213*, 104716.

Fox, W. F., Luna, L., & Schaur, G. (2014). Destination taxation and evasion: Evidence from us inter-state commodity flows. *Journal of Accounting and Economics*, *57*(1), 43–57.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, *225*(2), 254–277.

Table 4.8: Frozen Treatment Results (Log-Lin Specification, Staggered Adoption Check)

	Log(Firms) (1)	Log(Estabs) (2)	Log(Emp) (3)	Log(BA) (4)
Threshold (Frozen)	−0.0001*	−0.0001*	−0.0002**	−0.0008
	(0.0001)	(0.0001)	(0.0001)	(0.0005)
State & Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	557	557	557	557
Within R^2	0.729	0.763	0.750	0.228

Notes: Panel regressions with state and year fixed effects. Standard errors clustered at the state level in parentheses. The treatment variable is constructed by freezing each state's threshold at its first observed positive value, held constant from 2019 through 2023; all states are assigned zero exposure prior to 2019, removing variation in adoption timing across states. Controls include: population, median income, top 5% income, college degree %, non-white %, unemployment rate, service sector share, government spending per capita, property crime rates, high-speed internet penetration, max income tax rate, max corporate tax rate, coal production, and oil production. *p<0.1; **p<0.05; ***p<0.01

Vita

Lucas Gurgel was born in Fortaleza, Brazil. He earned a Bachelor's degree in Electrical Engineering at the Federal University of Ceará, where he later served as an undergraduate teaching assistant in mathematics and pre-engineering courses. After graduation, he began a career as an electrical engineer, working in wind farm operations and university infrastructure projects in Brazil. Motivated by a belief that individual choices and policy decisions shape economic outcomes, he transitioned to economics and earned a Master's degree in Economics from the Federal University of Ceará. In 2022, he joined the University of Tennessee to pursue a doctoral degree in Economics, specializing in Public Economics and Labor Economics. His dissertation investigates the effects of broadband expansion on human capital accumulation and business activity, with a particular focus on how high-speed internet connectivity shapes ADHD outcomes and the demand for treatment across the United States. As a Graduate Research Assistant at the Boyd Center for Business and Economic Research, Lucas develops policy evaluation projects in the economics of education for faculty and external organizations. He is a dedicated educator committed to using empirical evidence to improve public policy outcomes.