NSF CSSI Elements Proposal

Scalable Policy Microsimulation Infrastructure

PI: Max Ghenis, PolicyEngine / PSL Foundation **Co-PI**: Nikhil Woodruff, PolicyEngine
Requested Amount: $599,968 over 3 years

Project Summary

Overview

Public economics research increasingly requires modeling the joint tax-benefit system—how income taxes, payroll taxes, and benefit programs like SNAP, Medicaid, and the EITC interact to shape household incentives and outcomes. Yet existing cyberinfrastructure was designed for an earlier era: TAXSIM models income taxes but not benefits; TRIM3 models benefits but is only accessible through commissioned reports; neither supports independent reform modeling. Researchers studying effective marginal tax rates, benefit cliffs, or optimal policy design must cobble together separate tools or build custom solutions.

PolicyEngine is open-source infrastructure that models the complete U.S. tax-benefit system—federal and state income taxes, payroll taxes, and major benefit programs—in a unified framework. Users can calculate current-law outcomes or model any hypothetical reform. Everything runs on calibrated microdata via modern Python APIs, and every calculation traces to 1,800+ statutory citations embedded in the codebase. We are validated against NBER's TAXSIM (MOU with NBER; Dan Feenberg serves as advisor) and the Atlanta Fed Policy Rules Database. Current users include the Joint Economic Committee, UK Treasury, and researchers at USC, Harvard, and Georgetown.

This project will modernize PolicyEngine's core infrastructure to support population-scale research. We will build **continuous validation infrastructure** that automatically compares every release against TAXSIM and Atlanta Fed benchmarks, ensuring accuracy as law changes annually. We will deliver **performance improvements** enabling microsimulation across 150+ million tax units for distributional research and policy optimization. We will create **R and Stata interfaces** to meet economists in their preferred environments. And we will optimize **memory-efficient branching** for complex tax provisions and provide **cloud research infrastructure** enabling population-scale analysis without high-end local hardware.

The research opportunity is substantial. Consider a researcher studying how proposed CTC expansions interact with SNAP phase-outs to affect labor supply—this requires modeling taxes and benefits together with reform capability at population scale. Or consider studying effective marginal tax rates across states—this requires 50-state tax and benefit models in a unified framework. PolicyEngine enables this research, but core infrastructure modernization is required to support it reliably at scale.

Intellectual Merit

This project creates open-source cyberinfrastructure enabling research on the joint tax-benefit system at population scale. The intellectual contributions include: (1) continuous cross-model

validation methodology establishing accuracy benchmarks for microsimulation; (2) performance architecture enabling real-time policy optimization over population samples; (3) technical approaches for modeling complex tax provisions with scenario branching; and (4) reproducible infrastructure where every calculation traces to authoritative sources.

Most significantly, the infrastructure enables research questions currently intractable with existing tools. How do SNAP benefit reductions interact with EITC phase-outs to affect labor supply? What policy combinations minimize poverty while maintaining work incentives? How do effective marginal tax rates vary across states due to benefit program differences? These questions require modeling taxes and benefits together at population scaleâ€"precisely what PolicyEngine provides.

Broader Impacts

Open tax-benefit infrastructure democratizes policy analysis capabilities currently restricted to well-resourced institutions. Graduate students can conduct dissertation research without proprietary software licenses or restricted data access. State legislators can independently model policy alternatives rather than waiting weeks for external analysis. International researchers can study U.S. policy without institutional barriers.

The platform already demonstrates broader impact: partner applications have identified over $1 billion in unclaimed benefits across platforms including MyFriendBen, Amplifiâ€™s Benefit Navigator, Student Basic Needs Coalition, Mirza, and Starlight. Every calculation is version-controlled and traceable to statute, enabling reproducible research. Commercial applications in tax preparation and financial planning provide sustainability pathways ensuring long-term maintenance without perpetual grant funding.

The goal is cyberinfrastructure as foundational for tax-benefit research as TAXSIM has been for tax research aloneâ€"but open, comprehensive, and designed for modern research workflows.

Keywords

Tax-benefit microsimulation; open-source infrastructure; policy reform modeling; TAXSIM; effective marginal tax rates; SNAP; EITC; computational economics; reproducible research; rules as code

---

Project Description

1. The Research Infrastructure Gap

1.1 The Need for Joint Tax-Benefit Modeling

Modern welfare states operate through two interacting systems: taxes that reduce disposable income and benefits that increase it. A single mother earning $25,000 might receive SNAP benefits, qualify for Medicaid, claim the EITC and CTC, and owe payroll taxesâ€"all simultaneously. Whether she should work additional hours depends on how all these programs interact: the EITC phase-in increases incentives, but SNAP benefit reduction and potential Medicaid loss (â€œthe cliffâ€�) decrease them.

Yet research infrastructure developed separately for taxes and benefits:

Tax-focused infrastructure (TAXSIM, tax microsimulation models) calculates income tax liabilities but ignores benefit programs. Decades of tax incidence research using these tools cannot capture how SNAP phase-outs or Medicaid eligibility affect the conclusions.

Benefit-focused infrastructure (TRIM3, state eligibility models) models benefit programs but typically ignores or simplifies tax treatment. Research on benefit adequacy or take-up rates misses tax interactions.

This fragmentation is not just inconvenient—it produces biased research. Studies of effective marginal tax rates that ignore benefits understate work disincentives for low-income households. Studies of benefit adequacy that ignore tax credits understate total support. Research on poverty measurement, labor supply, and optimal policy design all suffer when taxes and benefits are modeled separately.

1.2 Limitations of Existing Infrastructure

NBER TAXSIM: Created by Daniel Feenberg in 1993, TAXSIM has enabled foundational tax research with over 1,000 academic citations. But TAXSIM has fundamental limitations:

- *Taxes only*: No SNAP, Medicaid, TANF, SSI, or other benefit programs
- *Current law only*: Cannot model hypothetical reforms—only actual law for historical years
- *Closed source*: Researchers cannot inspect calculations or verify edge case handling
- *Individual records*: No integrated microsimulation on population samples
- *Legacy interface*: Web forms and batch files; no modern API

TRIM3 (Transfer Income Model): The Urban Institute's model includes both taxes and benefits but:

- *Not publicly available*: Analysis only through commissioned reports; model itself not accessible
- *No independent research*: Researchers cannot run their own analyses or test hypotheses
- *Not reproducible*: Proprietary code cannot be inspected or cited
- *No reform modeling*: Custom analysis requires contracting with Urban Institute

Similar limitations apply to other Urban Institute models (ATTIS for health, DynaSim for Social Security, HIPSM for health insurance)

CBO and JCT Models: Congressional budget offices have sophisticated internal models, but they are not public, institutional use only, and not available for academic research.

1.3 Research Questions Requiring Integrated Infrastructure

Consider what becomes possible with infrastructure modeling the joint tax-benefit system:

Effective Marginal Tax Rate Research: What is the true marginal tax rate facing households at different income levels? This requires summing income tax rates, payroll tax rates, SNAP benefit reduction rates, EITC phase-out rates, and potential Medicaid loss.

Tax-Benefit Interaction Studies: Decades of labor supply research estimates responses to income tax changes. But households respond to the full tax-benefit schedule. A study of EITC labor supply effects that ignores SNAP interactions may misattribute responses.

Optimal Policy Design: What combination of tax rates, credit parameters, and benefit levels minimizes poverty while maintaining work incentives? This is a constrained optimization problem intractable without integrated microsimulation.

State Policy Natural Experiments: The 50 states offer natural experiments in policy design. But research comparing state policies must model both state taxes and state benefit administration.

Poverty Measurement: The Supplemental Poverty Measure includes taxes and benefits in resources. Proper SPM analysis requires calculating both accurately.

---

2. PolicyEngine: Integrated Tax-Benefit Infrastructure

2.1 Technical Architecture

PolicyEngine is open-source infrastructure built on three core components:

Rules Engine (policyengine-core): A declarative policy specification language where tax and benefit rules are encoded as Python functions with embedded metadata.

Policy Models (policyengine-us, policyengine-uk, etc.): Country-specific implementations encoding actual law. PolicyEngine-US includes:

- Federal income taxes: All IRC provisions affecting individual filers
- State income taxes: Complete models for all 50 states plus DC
- Payroll taxes: Social Security, Medicare, unemployment insurance
- Federal benefits: SNAP, Medicaid, TANF, SSI, WIC, CHIP, Section 8, LIHEAP, Lifeline, school meals, CCDF
- Tax credits: EITC, CTC, CDCTC, AOTC, LLC, energy credits, state equivalents

Every calculation traces to authoritative sources through 1,800+ structured citations to the U.S. Code, CFR, and state statutes.

Data Infrastructure (policyengine-us-data): Enhanced microdata for population-scale analysis, calibrated to administrative totals with imputed variables.

2.2 Reform Modeling Capability

Unlike current-law-only tools, PolicyEngine enables hypothetical reform analysis:

```
from policyengine_us import Simulation

reform = {
    "gov.irs.credits.ctc.amount.base[0].amount": 3600,
    "gov.irs.credits.ctc.refundable.fully_refundable": True
}

simulation = Simulation(situation=household, reform=Reform.from_dict(refor
```

2.3 Demonstrated Adoption

Government Users: Joint Economic Committee, UK Cabinet Office/HM Treasury, NY State Senator Gounardes

Research Users: USC Center for Economic and Social Research, curriculum discussions at Berkeley, Georgetown, Northwestern, Harvard

Public Benefit: Partner applications identified $1B+ in unclaimed benefits (MyFriendBen $800M+ in Colorado, Amplifi $185M in California, plus Student Basic Needs Coalition, Mirza, Starlight)

Technical Scale: 146 GitHub repositories, 50,000+ commits, 9,034 encoded parameters, 30+ contributors, 620+ forks, NSF POSE Phase I award (#2229069)

2.4 Validation

- TAXSIM: Formal MOU with NBER; Dan Feenberg serves as advisor
- Atlanta Fed Policy Rules Database: Parameter cross-validation with feedback loop
- IRS Statistics of Income: Aggregate tax outcome validation
- Administrative benchmarks: Benefit enrollment and payment totals

---

3. Infrastructure Gaps Requiring NSF Investment

3.1 Continuous Validation Infrastructure

Problem: Current validation is reactive, incomplete, and unsustainable.

Solution: Automated validation comparing every code change against TAXSIM and PRD, with public accuracy dashboards and regression detection.

3.2 Performance for Population-Scale Research

Problem: Population-scale analysis (150M+ tax units) requires hours of computation.

Solution: Vectorization, parallelization, and memory optimization enabling 4x speedup.

3.3 R and Stata Interfaces

Problem: Economists predominantly work in R and Stata; Python-only access creates adoption barriers.

Solution: CRAN-installable R package and SSC-installable Stata package with native syntax.

3.4 Memory-Efficient Branching and Cloud Infrastructure

Problem: PolicyEngine supports branching (AMT, itemization elections), but memory requirements prevent full population simulations on standard hardware.

Solution: Memory optimization (smart caching, streaming computation) plus cloud research infrastructure enabling population-scale analysis.

---

4. Technical Approach

4.1 Continuous Validation Infrastructure

TAXSIM Pipeline: Automated test generation → CI integration → accuracy dashboard → regression detection

Atlanta Fed PRD: Parameter cross-reference → benefit program validation → bidirectional feedback

Targets: >98% federal TAXSIM agreement, >95% state agreement, >99% PRD parameter coverage

4.2 Performance Optimization

Strategy: Vectorization (NumPy where/select), parallelization (chunked processing), lazy evaluation

Targets:

Scenario	Current	Year 3 Target
Single household	50ms	20ms
100K households	3 min	45 sec
Full population (150M)	8 hours	90 min

4.3 R and Stata Interfaces

R Package: CRAN-compliant, tidyverse-compatible, vignettes for research workflows

Stata Package: SSC-installable, native syntax, survey integration

4.4 Memory Optimization and Cloud Infrastructure

Memory: Lazy branch evaluation, smart caching, streaming computation, checkpoint/restart

Cloud: PolicyEngine Cloud API, pre-configured cloud templates, workshop cloud access, hybrid workflows

---

5. Relationship to Existing NSF Investment

PolicyEngine is an active **NSF POSE Phase I awardee** (#2229069, through July 2026). POSE focuses on ecosystem development; CSSI focuses on technical infrastructure. **Non-overlapping and complementary.**

POSE Phase I (Ecosystem)	CSSI Elements (Infrastructure)
Community governance	Validation pipelines
Contributor onboarding	Performance optimization
I-Corps customer discovery	R/Stata interfaces
Documentation standards	Memory optimization & cloud

---

6. Timeline and Milestones

Year 1: TAXSIM CI validation (federal), R package alpha, accuracy dashboard beta

Year 2: 50-state TAXSIM validation, PRD validation, CRAN/SSC packages, 2x performance

Year 3: Memory optimization, cloud infrastructure, documentation suite, methodology paper

---

7. Broader Impacts

7.1 Democratizing Policy Analysis

- Graduate students: Publication-quality analysis without TRIM3 access
- State governments: Independent, immediate policy modeling
- International researchers: Remove institutional barriers

7.2 Reproducible Research

Version-controlled, traceable to statute, inspectable, testable

7.3 Research Training

Tutorial notebooks, workshop curriculum, video documentation

7.4 Public Benefit

Benefit navigation tools reaching 100,000+ individuals annually, identifying $1B+ in unclaimed benefits across MyFriendBen, Amplifi, and other partners

---

8. Team and Qualifications

Max Ghenis, PI: PolicyEngine CEO, NSF POSE Phase I lead, former Google data scientist

Nikhil Woodruff, Co-PI: PolicyEngine CTO, seconded to UK Cabinet Office/HM Treasury

Dan Feenberg, Advisor: TAXSIM creator, NBER, I-Corps mentor

John Sabelhaus, Advisor: Former Federal Reserve, Social Security modeling expert

---

9. Sustainability

Commercial Revenue: Tax preparation integration, financial planning tools, policy consulting

Government Contracts: State fiscal offices, federal agency integration

Community Maintenance: 30+ contributors, 620+ forks, POSE governance structures

---

Budget Summary

Category	Year 1	Year 2	Year 3	Total
Senior Personnel	$33,750	$33,750	$33,750	$101,250
Other Personnel	$95,333	$95,333	$95,334	$286,000
Fringe Benefits	$38,725	$38,725	$38,725	$116,175
Travel	$6,000	$6,000	$6,000	$18,000
Other Direct	$8,000	$8,000	$8,000	$24,000
Indirect	$18,181	$18,181	$18,181	$54,543
Total	**$199,989**	**$199,989**	**$199,990**	**$599,968**

References Cited

Feenberg, D., & Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1), 189-194.

Gruber, J., & Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of Public Economics*, 84(1), 1-32.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1), 3-50.

Chetty, R., Friedman, J. N., & Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review*, 103(7), 2683-2721.

Moffitt, R. (2002). Welfare programs and labor supply. *Handbook of Public Economics*, 4, 2393-2430.

Atlanta Federal Reserve Bank. (2024). Policy Rules Database. https://www.atlantafed.org/economic-mobility-and-resilience/advancing-careers-for-low-income-families/policy-rules-database

PolicyEngine. (2024). PolicyEngine-US Documentation. https://policyengine.github.io/policyengine-us/