ITC_Coq

Formal Verification of

Indirect Treatment Comparison Methods

A Comprehensive Analysis and Technical Report

Based on MathComp-Analysis

Coq 9.x / Rocq Prover

January 2026

Executive Summary

ITC_Coq is a formal verification project in Coq that formalizes the mathematical foundations of causal inference theory with a focus on Indirect Treatment Comparisons (ITC) used in healthcare technology assessments (HTA).

Key Accomplishments:
- 13 Coq files totaling approximately 4,941 lines
- Formalizes the Rubin Causal Model (potential outcomes framework)
- Proves the foundational Propensity Score Theorems (Rosenbaum & Rubin, 1983)
- Establishes Double Robustness of AIPW estimators
- Provides formal comparison of MAIC vs STC methods
- Built on MathComp-Analysis for measure-theoretic rigor

Codebase Overview

File Inventory

The project contains 13 Coq files organized in 4 phases:

Phase 1 - Foundations:
  Axioms.v (243 lines) - Classical logic foundations
  BasicTypes.v (347 lines) - Core type definitions
  PotentialOutcomes.v (439 lines) - Rubin Causal Model
  CausalAssumptions.v (529 lines) - Identifying assumptions

Phase 2 - Core Theory:
  ConditionalIndep.v (466 lines) - Measure-theoretic CI
  BalancingScore.v (364 lines) - Balancing score theory
  PropensityScore.v (441 lines) - Main PS theorems

Phase 3 - Estimators:
  IPWEstimator.v (304 lines) - IPW estimator
  OutcomeRegression.v (298 lines) - G-computation
  DoublyRobust.v (333 lines) - AIPW & DR property

Phase 4 - ITC Methods:
  MAIC.v (385 lines) - Matching-adjusted ITC
  STC.v (337 lines) - Simulated treatment comparison
  Comparison.v (455 lines) - MAIC vs STC analysis

TOTAL: ~4,941 lines

Key Theorems

1. Propensity Score Theorem (Rosenbaum & Rubin, 1983)

THEOREM: If treatment is strongly ignorable given X, then treatment is strongly ignorable given e(X).

If: (Y(0), Y(1)) independent of T | X and 0 < e(X) < 1
Then: (Y(0), Y(1)) independent of T | e(X)

SIGNIFICANCE: Enables dimension reduction - replace high-dimensional X with scalar e(X) without losing causal identification.

2. Double Robustness Theorem

THEOREM: AIPW is consistent if EITHER the propensity score OR the outcome model is correctly specified.

If: (e_hat = e_true) OR (mu_hat = mu_true)
Then: E[AIPW] = E[Y(1)]

SIGNIFICANCE: Provides "double protection" against model misspecification.

3. MAIC vs STC Comparison

The formal comparison establishes:

MAIC (Matching-Adjusted Indirect Comparison):
- Does NOT require outcome model
- REQUIRES good overlap
- Robust to model misspecification
- Variance inflates with poor overlap

STC (Simulated Treatment Comparison):
- REQUIRES correct outcome model
- Does NOT require overlap
- Can extrapolate to new populations
- Stable variance regardless of overlap

RECOMMENDATION: Use MAIC when overlap is good and outcome model uncertain. Use STC when overlap is poor and outcome model reliable.

Method Comparison

Comprehensive comparison of MAIC vs STC:

Criterion	MAIC	STC	Winner
Requires outcome model	No	Yes	MAIC
Requires good overlap	Yes	No	STC
Robust to model misspec	Yes	No	MAIC
Variance with poor overlap	HIGH	Normal	STC
Can extrapolate	No	Yes	STC
Non-collapsibility	Automatic	Manual	MAIC

Significance and Implications

Academic Significance

1. First formal verification of propensity score theory in Coq
2. Measure-theoretic foundations using MathComp-Analysis
3. Comprehensive coverage from basic definitions to advanced ITC methods
4. Educational value with extensive documentation

Practical Implications for HTA

1. Rigorous foundations for regulatory decision-making
2. Clear assumption statements required for any analysis
3. Method selection guidance formalized as decision rules
4. Quality assurance through machine-checked proofs

Compilation Environment

- Coq Version: 9.1.0 (Rocq Prover)
- Dependencies: coq-mathcomp-ssreflect >= 2.0
  coq-mathcomp-algebra
  coq-mathcomp-analysis >= 1.9.0
  coq-hierarchy-builder >= 1.6.0

References

1. Rosenbaum, P.R. & Rubin, D.B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." Biometrika, 70(1), 41-55.

2. Robins, J.M., Rotnitzky, A. & Zhao, L.P. (1994). "Estimation of Regression Coefficients When Some Regressors Are Not Always Observed." JASA, 89(427), 846-866.

3. Signorovitch, J.E. et al. (2010). "Comparative Effectiveness Without Head-to-Head Trials." PharmacoEconomics, 28(10), 935-945.

4. Phillippo, D.M. et al. (2018). "Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal." Medical Decision Making, 38(2), 200-211.

5. NICE TSD 18 (2016). "Methods for Population-Adjusted Indirect Comparisons in Submissions to NICE."

Report generated: January 2026
ITC_Coq Project