The U.S. Health Care System and Single-Payer Reform: An Evidence Review

A comprehensive, evidence-based review of the published literature comparing the current U.S. health care system with universal-coverage and single-payer models

Prepared August 2026

Preface and Methodology

This review was commissioned to answer a specific question as neutrally as the evidence allows: what does the published literature actually tell us about the likely implications of adopting a single-payer health care system in the United States? It does not begin from the premise that single payer is either desirable or undesirable, and it treats institutional reputation — on any side of this debate — as a reason for scrutiny, not a substitute for it. Every quantitative claim in this review is sourced to a specific, named study, report, or dataset, verified via direct retrieval wherever possible; where a source could not be independently verified during this review's research process, that limitation is stated explicitly rather than glossed over, and no citation appears in this document for a source whose existence could not be confirmed. Two instances of this are worth flagging at the outset because they illustrate the review's verification standard in practice: an oft-referenced "GAO report comparing roughly 20 single-payer cost studies" could not be located and does not appear to exist as described (the closest real analogue is a 2020 systematic review published in PLOS Medicine, cited throughout as such); and a widely repeated claim that RAND analyzed Colorado's 2016 single-payer ballot initiative is incorrect — that analysis was conducted by the Colorado Health Institute, a different organization, and is cited accordingly.

The review draws on government analyses (the Congressional Budget Office, the Government Accountability Office, CMS's Office of the Actuary), nonpartisan research institutions (RAND, the Urban Institute, the Commonwealth Fund, KFF), peer-reviewed academic literature across health economics and health services research, and analyses produced by organizations with explicit institutional or ideological commitments on both sides of the single-payer debate — including single-payer advocacy (Physicians for a National Health Program) and organized opposition (the Heritage Foundation, the Mercatus Center, the Partnership for America's Health Care Future, an industry coalition). Every source with an institutional or advocacy affiliation is flagged as such at first mention and in the annotated bibliography, consistent with the instruction not to treat institutional reputation as a proxy for correctness in either direction. Sources current through August 2026 were used; where a cited study or projection predates significant subsequent policy changes (the 2022 Inflation Reduction Act's Medicare drug-negotiation provisions, the 2025 federal reconciliation law's coverage effects), this is noted.

Two organizing distinctions run throughout this review and are essential to reading it correctly. First, "single-payer financing" as an abstract mechanism (one public payer collecting funds and paying providers, as in Canada or Taiwan) is analytically distinct from the specific, more comprehensive Medicare for All legislative proposals debated in U.S. politics, which typically bundle single-payer financing with a much richer benefit package and near-total elimination of cost-sharing than most existing single-payer systems abroad actually provide. Evidence about one does not automatically transfer to the other, and this review is explicit throughout about which is being discussed. Second, cross-national comparisons of financing structure and cost (administrative overhead, drug prices, negotiated rates) are far more directly interpretable than cross-national comparisons of health outcomes (life expectancy, mortality), which are jointly determined by financing, demographics, behavior, and social conditions that vary enormously across the countries typically compared to the United States. This review applies that distinction consistently rather than treating all cross-national comparisons as equally probative of the financing question specifically.

The review is organized into twelve sections following the structure requested: a taxonomy of health-system models; a systematic synthesis of the cost and outcomes literature; identification of the major analyses; a substantial annotated bibliography; a critical, assumption-by-assumption evaluation of the major cost models; a explicit sorting of findings by confidence level; an examination of the mechanisms behind international cost and outcome differences; a direct evaluation of single-payer-specific implications; a comparison of competing policy scenarios; an analysis of why credible analysts disagree; an evidence hierarchy; and a final synthesis, including the requested 1,000–1,500 word bottom-line conclusion.

1. Defining the Comparison: What "Single Payer" Does and Does Not Mean

Public debate over "single payer" routinely collapses several structurally distinct models into one category, which is one reason the evidence surrounding it is so often talked past rather than engaged with. Before any evidence can be weighed, the terms need to be pinned down using the standard health-systems typology employed by the Commonwealth Fund's International Health Policy Center and traceable to Milton Roemer's classic health-systems classification work (a framework later popularized, including by single-payer advocacy groups such as Physicians for a National Health Program, but whose underlying substance is independently corroborated by Commonwealth Fund's country-by-country system profiles).

The current U.S. multipayer system. A patchwork of employer-sponsored private insurance (≈49% of the population), Medicare (a federal single-payer program for those 65+ and some disabled individuals), Medicaid (a joint federal-state program for lower-income individuals, administered separately by 50 states), the ACA marketplaces (regulated, subsidized private insurance), the VA and Indian Health Service (a government-run direct-delivery model within the larger system), and roughly 26–27 million uninsured people. No other high-income country combines this many financing mechanisms, each with its own eligibility rules, provider networks, and payment rates, inside one national system.

National Health Service (NHS) / "Beveridge" models. General taxation funds a government-owned and largely government-operated delivery system; hospitals and, in some cases, physicians are direct government employees. The United Kingdom (1948) and Italy are the standard examples. This model consolidates financing *and* delivery under public ownership — a more thoroughgoing form of government involvement than single-payer financing alone.

Social Health Insurance (SHI) / "Bismarck" models. Financing runs through multiple, tightly regulated, competing non-profit "sickness funds" financed by payroll contributions, with care delivered by a mix of private and public providers under uniform, often nationally negotiated, fee schedules. Germany and, since reforms extending coverage to nearly all residents, France are the standard examples. Providers remain largely private; insurers remain plural; but government-set or all-payer-negotiated rates eliminate most of the price variation and administrative fragmentation associated with the U.S. system.

National Health Insurance (NHI) / single-payer models. One public (or quasi-public) insurer collects funds (via taxes, payroll contributions, or premiums) and pays private and public providers directly; delivery itself remains substantially private. Canada (provincially administered, publicly funded) and Taiwan (which in 1995 merged several prior public insurance schemes into a single national plan) are the standard examples. This is the model most often meant by "single payer" in the U.S. debate, and it is financing consolidation without delivery-system nationalization — a distinction frequently lost in political shorthand that conflates single payer with "socialized medicine" (an NHS-style model).

Regulated multipayer systems with private insurers. Universal coverage is achieved by mandating that all residents purchase tightly regulated private insurance, with community rating, guaranteed issue, and risk adjustment eliminating medical underwriting. The Netherlands and Switzerland are the standard examples. This model demonstrates that universal coverage and non-single-payer financing are not mutually exclusive — a fact directly relevant to any claim that only single payer can achieve universality.

Medicare for All proposals specific to the United States. These are not a single model but a family of legislative proposals (the most prominent being the Medicare for All Act, S.1804/H.R.1384 in various Congresses) that typically combine (a) single-payer financing, (b) a comprehensive benefit package often including long-term services and supports, dental, vision, and hearing, (c) elimination or near-elimination of patient cost-sharing, and (d) abolition of the role of private insurers for covered benefits. This bundling matters enormously for the evidence review that follows: most quantitative cost estimates are estimates of *a specific bundled U.S. legislative proposal*, not of "single-payer financing" as an abstract mechanism. A country can adopt single-payer financing without also adopting zero cost-sharing, comprehensive long-term-care coverage, or the complete exclusion of private insurance (Canada's system, for instance, retains a substantial private role for drugs, dental, and vision, and permits modest cost-sharing in some provinces for services outside the core "medically necessary" bundle). Evidence about Canada's single-payer system, therefore, does not automatically transfer to evidence about the Medicare for All Act of 2017, and vice versa — a distinction this review maintains throughout, and one that a substantial share of the political debate on both sides collapses.

Where comparisons are genuinely comparable — and where they are not

International comparisons of *financing mechanism and cost structure* (administrative overhead, drug prices, negotiated provider rates) are reasonably comparable across countries, because these are structural, measurable features of how each system organizes payment. International comparisons of *health outcomes* are far less straightforward, because outcomes are jointly determined by the financing system, the delivery system, population demographics, behavioral risk factors (obesity, smoking, opioid and firearm mortality), social determinants of health (poverty, inequality, housing), and public health infrastructure — factors that vary enormously across the countries typically used as U.S. comparators and that are not attributable to the financing model itself. The National Academies' 2013 panel report *U.S. Health in International Perspective* is explicit on this point, and its caution is treated as a standing methodological rule throughout this review: findings about spending and prices are more directly interpretable as "caused by" system design than findings about mortality and life expectancy, which require much more caution before attributing them to the financing model rather than to the confounders above.

Comparisons of specific state-level or bill-specific U.S. proposals (Vermont, Colorado, New York, California) to each other are directly comparable, since they were commissioned in comparable contexts using overlapping methodologies (often the same modeling teams). Comparisons of these state proposals to national single-payer countries are far less comparable, because a single U.S. state cannot capture population risk pools the way a national system can, cannot set its own reimbursement floor for federal programs (Medicare, ERISA-governed self-insured employer plans) without a federal waiver, and faces adverse-selection and out-migration risks that a national system does not. This is a major reason why Vermont's 2014 abandonment and Colorado's 2016 ballot defeat are informative about the *specific fiscal and political constraints facing single-state implementation*, but only weakly informative about the fiscal feasibility of a *national* single-payer system, which would not face the same interstate boundary problems.

2. Systematic Literature Review

2.1 Cost

Total national health expenditure. U.S. national health expenditure (NHE) reached $5.3 trillion in 2024 — $15,474 per person, 18.0% of GDP — per CMS's Office of the Actuary, the standard official data series (CMS, "National Health Expenditure Fact Sheet," 2024 data). CMS's June 2026 projections put 2025 spending at roughly $5.7 trillion and project NHE will reach approximately $9.0 trillion, 20.6% of GDP, by 2034, with growth driven initially by drug demand and post-pandemic utilization and later tempered by ACA marketplace and Medicaid coverage losses tied to 2025 federal legislation. No other high-income country spends close to 18% of

GDP on health care; the OECD median is roughly 11%, and Switzerland, the second-highest spender, is still several points of GDP below the U.S.

Public versus private spending. Public spending (federal plus state/local) accounts for approximately 48% of U.S. NHE, private spending (employer, household, and other private sources) for approximately 52% (CMS NHE Fact Sheet, 2024; Peterson-KFF Health System Tracker). This split is frequently misunderstood in political debate: the U.S. government already directly finances nearly half of all health spending — comparable in absolute terms to what many single-payer countries spend in total — even before any expansion of public coverage.

Employer and household costs. The 2025 KFF Employer Health Benefits Survey found average annual premiums of $9,325 for single coverage and $26,993 for family coverage, with average worker contributions of $1,440 and $6,850 respectively. Thirty-four percent of covered workers face a single-coverage deductible of $2,000 or more. Out-of-pocket spending reached $556.6 billion in 2024 (11% of NHE). Medical debt affects a substantial share of the population: KFF estimates $220 billion in outstanding U.S. medical debt, and the Commonwealth Fund's 2024 biennial survey found 23% of insured adults are effectively underinsured (facing high enough cost-sharing relative to income that coverage provides only partial financial protection).

Prescription drug spending. RAND's most recent international drug price comparison (2024 report, 2022 data) found U.S. manufacturer-level drug prices were 278% of the average across 33 comparator OECD countries overall, and 422% of the comparator average for brand-name drugs specifically. Generic drugs, which make up 90% of U.S. prescription volume, are actually priced *below* the comparator average — the entire multiplier is a brand-name-drug phenomenon, concentrated in the roughly 10% of prescriptions that account for 87% of U.S. drug spending.

Hospital and physician spending. Hospital care ($1.63 trillion, ~31% of NHE) and physician/clinical services ($1.11 trillion, ~21% of NHE) together account for roughly half of all U.S. health spending (CMS NHE Fact Sheet, 2024).

Administrative costs. This is one of the best-documented and most consequential findings in the entire literature. Multiple independently conducted, methodologically distinct studies converge on the same qualitative conclusion, even though they differ substantially in the size of the estimate: U.S. health-system administrative costs are dramatically higher, in both absolute and relative terms, than in any peer country. Himmelstein, Woolhandler, and colleagues (Annals of Internal Medicine, 2020) estimated total U.S. health-administration spending at $812 billion in 2017 — 34.2% of measured national health expenditures — versus 17.0% in Canada, a per-capita gap exceeding $1,900. Their earlier work (NEJM 2003; Health Affairs 2014) found similar multi-fold administrative-cost gaps specifically in hospital administration (U.S. 25.3% of hospital spending versus roughly 12% in Canada and Scotland) and insurance overhead. Independently, Papanicolas, Woskie, and Jha (JAMA 2018), using a completely different data source and research team not affiliated with single-payer advocacy, found administrative costs of running the health system consumed 8% of total U.S. health expenditure versus roughly 3% across ten other high-income comparator countries — a smaller multiplier than Himmelstein and Woolhandler's estimate but the same directional and qualitative conclusion. A micro-costing study at a single U.S. academic medical center (Tseng et al., JAMA 2018) found physician billing and insurance-related administrative costs ranging from $20 to $215 per patient encounter, consuming up to a quarter of professional revenue for some encounter types — direct, bottom-up confirmation that U.S. billing complexity carries a real and substantial cost, independent of the top-down national-accounting studies. It should be noted that the size of the Himmelstein/Woolhandler estimates specifically has been formally disputed in the same journals that published them (a companion NEJM editorial explicitly challenged their 2003 methodology), and that their institutional affiliation with PNHP, an advocacy organization, is a relevant consideration in weighing how conservative or aggressive their assumptions are — a point elaborated in Section 5. What survives this dispute, however, is the qualitative finding, replicated by non-advocacy-affiliated researchers using different methods: U.S. administrative overhead is genuinely, substantially higher than in peer countries, by a factor of somewhere between roughly 2.5 and 4, depending on methodology.

Potential administrative savings under single payer, and the limits of that evidence. The Congressional Budget Office's own methodology papers (Working Paper 2020-08, 2020; Working Paper 2022-02, 2022) estimate that a single-payer design could reduce administrative spending to roughly 1.5–2% of total spending, versus roughly 12% for current private insurers, producing an estimated $533–562 billion in administrative savings by 2030 (about 1.8% of GDP) under CBO's most detailed illustrative model. Critically, CBO's own analysis shows this administrative saving does *not* automatically translate into lower total national health spending, because it can be more than offset by increased utilization (induced demand from expanded coverage and reduced cost-sharing) and by the specific provider-payment-rate assumptions built into any given design — CBO's five illustrative scenarios span a range from a $743 billion *decrease* to a $290 billion *increase* in total national health spending by 2030, depending almost entirely on those two additional variables, not on the size of the administrative saving itself (which is comparatively stable across the five scenarios). This is among the most important findings in the entire literature for resolving the political debate: the disagreement over whether single payer saves money is overwhelmingly a disagreement about provider payment rates and induced utilization, not about whether administrative savings exist (nearly every credible source, across the ideological spectrum, agrees some administrative savings would occur) or how large they are in isolation.

2.2 Outcomes

Life expectancy. U.S. life expectancy was 78.4 years in 2023 (CDC/NCHS), roughly 3.7 years below the average of the ten comparator nations used in Commonwealth Fund's *Mirror, Mirror* series. KFF's decomposition of this gap attributes roughly 32% to chronic disease, 24% to COVID-19, and 12% to substance use, with U.S. homicide and traffic-fatality rates (roughly eight and three times the peer average respectively) contributing further. This decomposition matters directly for the causal question at the heart of this review: a large share of the U.S. life-expectancy shortfall is driven by causes of death (homicide, drug overdose, traffic fatalities, obesity-related chronic disease) that are not primarily a function of health-insurance financing and would not be mechanically fixed by adopting a different payment system.

Avoidable/amenable mortality. This is the outcome measure most directly connected to health-*care* system performance specifically (as opposed to broader social and behavioral determinants), since it is constructed from causes of death considered preventable with timely, effective medical care. The foundational methodology (Nolte & McKee, Health Affairs 2008) ranked the U.S. last of 19 comparator countries. The Commonwealth Fund's 2024 *Mirror, Mirror* update again ranks the U.S. last of ten high-income nations overall, and last specifically on health outcomes and equity, while ranking second on "care process" measures (e.g., preventive-care and safety-related process measures) — a genuinely important nuance: the U.S. system is not uniformly poor across every dimension Commonwealth Fund measures, and in fact performs comparatively well on some process-of-care indicators even while ranking last on outcomes and equity.

Infant and maternal mortality. U.S. maternal mortality (18.6 per 100,000, per Commonwealth Fund's 2025 analysis) is roughly 3.6 times the peer-country average (5.1 per 100,000, per KFF). The within-U.S. racial disparity in this measure, however, is larger than the U.S.-versus-peer-country gap: Black maternal mortality in the U.S. (50.3 per 100,000) is comparable to rates in Vietnam or Brazil — a finding with direct relevance to the equity dimension of this review, since it suggests the U.S. maternal-mortality problem is not evenly distributed and that a substantial part of the U.S. disadvantage is a distributional/access problem within the current system rather than a uniform system-wide deficiency.

Cancer survival. By contrast, the CONCORD-3 global cancer-survival study (Allemani et al., Lancet 2018) found the U.S. among the best-performing countries in the world on cancer-specific five-year survival — roughly 90% for breast cancer and effectively at ceiling for prostate cancer. This is a genuine bright spot for the current U.S. system and a useful corrective to any claim that the U.S. underperforms peer countries uniformly across all clinical outcomes; it does not, however, resolve questions about the substantial share of the U.S. population that lacks timely access to that same high-quality cancer care in the first place (see "unmet need," below).

Cardiovascular outcomes and patient safety. The U.S. also outperforms the peer-country average on 30-day mortality following acute myocardial infarction and stroke, and performs well on post-surgical pulmonary embolism rates — further evidence that clinical quality, where care is actually delivered, is not uniformly worse in the U.S. system.

Unmet medical need and financial hardship. Here the U.S. disadvantage is stark and well-documented. Commonwealth Fund's International Health Policy Survey (2023) found nearly half of lower- and average-income U.S. adults report skipping needed care because of cost — far above any peer country in the survey. This is one of the clearest, most direct, and least confounded findings in the entire outcomes literature, because it measures access and financial burden directly rather than inferring them from downstream mortality statistics that are also affected by non-financial factors.

Health disparities within the United States. Commonwealth Fund's 2026 State Health Disparities Report finds that within-U.S. racial and income-based disparities are often as large as, or larger than, the gap between the U.S. as a whole and peer countries — a finding with direct implications for how much of the "U.S. versus other countries" framing obscures a more consequential "U.S. advantaged groups versus U.S. disadvantaged groups" reality.

Causal versus correlational evidence. Most cross-national outcome comparisons above are correlational: they compare aggregate national statistics without controlling for the demographic, behavioral, and socioeconomic confounders that plausibly explain a substantial share of the gap (see Section 7). The strongest causal evidence in the domestic literature on the specific question of whether expanding insurance coverage improves health comes from the Oregon Health Insurance Experiment — a true randomized controlled trial created when Oregon used a lottery to allocate a limited number of Medicaid slots (Finkelstein, Baicker, and colleagues, NEJM 2013 and companion QJE papers). It found clear, causally identified improvements in financial protection and mental health (reduced depression) among the newly insured, but was statistically underpowered to detect effects on physical-health biomarkers (blood pressure, cholesterol, glycemic control) or mortality within its study window — a limitation frequently overlooked in political debate, where the null finding on physical-health biomarkers is sometimes cited as if it were strong evidence that "insurance doesn't improve health," when the correct interpretation is that the study could not detect an effect of the size it would have needed to detect, not that it detected zero effect. Quasi-experimental studies of the ACA's Medicaid expansion (Sommers and colleagues; Miller, Johnson, and Wherry, NBER Working Paper 26081) using comparison-state designs generally find mortality reductions associated with expansion, though these designs are more vulnerable to confounding than a true randomized trial. A quasi-experimental study of Massachusetts's 2006 reform (Sommers, Long, and Baicker, Annals of Internal Medicine 2014) found a 2.9% relative reduction in all-cause mortality following the reform, though this finding has been subject to a published statistical critique questioning the robustness of its county-comparison design, and should be treated as moderate- rather than high-confidence. Taiwan's 1995 implementation of national health insurance offers a particularly clean natural experiment: a joinpoint-regression study found the decline in *amenable* mortality accelerated sharply immediately after the reform, while *non-amenable* mortality (deaths not considered preventable by medical care) showed no corresponding shift — precisely the pattern one would expect if the coverage expansion, specifically, drove the amenable-mortality improvement, since only the outcome category the reform could plausibly affect actually moved.

2.3 Synthesis of the cost and outcomes literature

Two conclusions can be stated with high confidence based on the above: first, the United States spends far more, largely because of higher prices for the same underlying units of care (drugs, physician services, hospital procedures) and higher administrative overhead, not because of higher utilization — U.S. utilization is actually below the peer-country average on many measures. Second, the U.S. outcome disadvantage is real but highly uneven: the U.S. underperforms peer countries substantially on some measures most directly tied to system access and equity (unmet need, maternal mortality, amenable mortality, financial hardship) while matching or outperforming them on others most directly tied to clinical quality once care is actually delivered (cancer

survival, acute cardiovascular mortality, patient safety). This unevenness is itself an important finding: it argues against any simple narrative in either direction — neither "the U.S. health care system delivers poor-quality care" nor "the U.S. system's high spending buys uniformly superior outcomes" is well supported by the evidence; the more defensible characterization is that the U.S. system delivers comparatively good clinical care to those who can access it, at a comparatively high price, while leaving a comparatively large share of the population with inadequate or delayed access — and that this access gap, not a clinical-quality gap, is the primary channel through which financing-system design plausibly affects health outcomes.

3. The Major Analyses Identified for This Review

Consistent with the research standards requested, the analyses drawn on most heavily in this review — and profiled in the annotated bibliography in Section 4 — are:

Government/nonpartisan analytic bodies:

- Congressional Budget Office: "Key Design Components and Considerations for Establishing a Single-Payer Health Care System" (2019); "How CBO Analyzes the Costs of Proposals for Single-Payer Health Care Systems" (Working Paper 2020-08); "Economic Effects of Five Illustrative Single-Payer Health Care Systems" (Working Paper 2022-02); "A Public Option for Health Insurance in the Nongroup Marketplaces" (2021); "Budgetary Effects of a Policy That Would Lower the Age of Eligibility for Medicare to 60" (2022); most recent baseline coverage projections (Feb. 2026).
- Government Accountability Office: "Canadian Health Insurance: Lessons for the United States" (GAO/HRD-91-90, 1991) and its 1992 follow-on costing report; "Highlights of a Forum: Reducing Spending and Enhancing Value in the U.S. Health Care System" (GAO-25-107465, 2025).
- CMS Office of the Actuary: National Health Expenditure Accounts and 2025–34 projections.
- National Academies of Sciences, Engineering, and Medicine / National Academy of Medicine: *U.S. Health in International Perspective: Shorter Lives, Poorer Health* (2013); *Vital Directions for Health & Health Care* (2017).

Research institutions generally regarded as nonpartisan or center-independent:

- RAND Corporation: international drug price comparisons (2021, 2024); New York Health Act assessment (2018); national Medicare for All spending estimate (2019/2020); public option analysis (2020).
- Urban Institute: cost estimates of the Sanders single-payer plan (2016, 2018); public option and Medicare buy-in analyses (2021–2022).
- Commonwealth Fund: *Mirror, Mirror* international rankings (2024); international health system profiles; state health disparities reporting.
- KFF (Kaiser Family Foundation): Employer Health Benefits Survey; uninsured-rate tracking; medical-debt analysis.
- Peterson-KFF Health System Tracker: cross-national spending, price, and utilization comparisons.
- OECD and WHO: comparative health statistics (cited via secondary compilations in this review; primary OECD Health at a Glance tables were not independently re-pulled — see limitations noted in Section 4).

Peer-reviewed academic literature:

- Himmelstein, Woolhandler, and colleagues (NEJM, Health Affairs, Annals of Internal Medicine): administrative-cost comparisons between the U.S. and Canada.
- Papanicolas, Woskie, and Jha (JAMA 2018): the most-cited modern cross-national spending-driver analysis.

- Anderson, Reinhardt, Hussey, and Petrosyan (Health Affairs 2003, 2019): the foundational "It's the Prices, Stupid" studies.
- Cooper, Craig, Gaynor, and Van Reenen (QJE 2019): hospital consolidation and pricing.
- Cai, Runte, Kahn, and colleagues (PLOS Medicine 2020): systematic review of 22 single-payer cost models.
- Galvani and colleagues (Lancet 2020): Yale model of the Medicare for All Act.
- Mello, Chandra, Gawande, and Studdert (Health Affairs 2010): the definitive costing of the U.S. medical liability system.
- Finkelstein, Baicker, and colleagues: the Oregon Health Insurance Experiment.
- Sommers and colleagues: Massachusetts and Medicaid-expansion mortality studies.

Analyses associated with a specific advocacy or institutional perspective (flagged explicitly wherever cited):

- Physicians for a National Health Program (PNHP) — single-payer advocacy; several of its co-founders' academic papers are cited above and are treated as *both* peer-reviewed *and* advocacy-affiliated.
- Mercatus Center (Charles Blahous) — free-market-oriented, George Mason University-housed, substantially associated with Koch-network funding.
- The Heritage Foundation, the Manhattan Institute, and the Committee for a Responsible Federal Budget — conservative or deficit-hawk-oriented critiques of Medicare for All costing.
- The Partnership for America's Health Care Future — an industry coalition (hospitals, insurers, pharmaceutical manufacturers) formed explicitly to oppose single-payer and Medicare expansion.
- Political Economy Research Institute (PERI), UMass Amherst — produced the costing used by the 2018 Sanders Medicare for All plan.

Consistent with the instruction not to treat institutional reputation as a proxy for correctness, each of these sources is evaluated on its stated methodology and assumptions in Sections 4 and 5, not on the credibility of its sponsoring institution alone. As will become clear, the most useful single fact to emerge from comparing these sources is that the *ideologically opposed* Mercatus Center and Urban Institute reached strikingly similar ten-year federal cost estimates for the 2016–2018 Sanders-style Medicare for All proposal (roughly $32–33 trillion) — a convergence that Blahous himself has pointed to publicly — while the estimates that diverge sharply from that range (PERI, RAND's national model, the Galvani/Yale model) diverge primarily because they model a materially different, generally less comprehensive, benefit and payment-rate design, not because of a straightforward left-right split in methodology.

4. Annotated Bibliography

Sources are ordered roughly by topic cluster (cost/administration, cost models of single payer, outcomes, price mechanisms, and critiques), not strictly by importance. Primary research and government/institutional primary sources are prioritized over secondary reporting. Every entry below was independently verified via direct retrieval of the source or its official abstract/landing page; URLs and identifiers (DOI/PMID/report number) are provided where available.

---

1. CMS Office of the Actuary — National Health Expenditure Accounts

Citation: Centers for Medicare & Medicaid Services, Office of the Actuary. "National Health Expenditure Data: NHE Fact Sheet" (2024 data, published 2025) and "National Health Expenditure Projections, 2025–34" (2026).

Year: 2025–2026 (annual series). Institution: U.S. federal government (CMS).

Type: Official administrative/actuarial data compilation and projection model.

Research question: What did the U.S. spend on health care, by category and source of funds, and what is it projected to spend?

Data sources: Comprehensive national accounts drawing on insurer, provider, and government program administrative data.

Methodology: National-accounts-style aggregation; projections use an actuarial macroeconomic model incorporating enrollment, price, and utilization trend assumptions.

Principal findings: $5.3 trillion (18.0% of GDP) in 2024; projected to reach ~$9.0 trillion (20.6% of GDP) by 2034; public/private split roughly 48/52.

Major assumptions: Projections assume continuation of current law and enacted 2025 legislative changes; sensitive to unanticipated policy shifts.

Strengths: Authoritative, methodologically transparent, the standard reference series used by virtually every other study in this review.

Limitations: Projections, by nature, carry uncertainty that grows with time horizon; the underlying "national health expenditure" definition includes some administrative and public-health spending categories that differ slightly from other countries' national accounts, complicating exact cross-national comparison (a caveat OECD/Commonwealth Fund comparative work generally adjusts for).

Conflicts of interest: None identifiable; production is a statutory government function insulated from electoral politics, though its projections can be affected by the policy assumptions built in by the administration in power.

Evidentiary support for conclusions: Data, not argued conclusions; the appropriate standard is data-quality, which is high.

Use in policy debate: The near-universal starting baseline for every cost estimate in this review, cited by RAND, CBO, Urban Institute, Mercatus, and PNHP alike.

Reliability assessment: Very high for historical data; moderate-high for near-term projections; appropriately treated as a fact base rather than an argument.

---

2. RAND Corporation — International Prescription Drug Price Comparisons: Estimates Using 2022 Data

Citation: Mulcahy AW, Schwam D, Lovejoy SL. RAND Research Report RR-A788-3 (Feb. 2024), prepared for HHS/ASPE. https://www.rand.org/pubs/research_reports/RRA788-3.html

Year: 2024 (2022 data). Institution: RAND Corporation (federally commissioned).

Type: Cross-national price comparison, quantitative.

Research question: How do U.S. manufacturer-level prescription drug prices compare to those in other high-income countries?

Data sources: IQVIA MIDAS multinational sales and volume data.

Methodology: ATC-code product matching, purchasing-power-adjusted price ratios, US-utilization-weighted averages, sensitivity analysis by market segment.

Principal findings: U.S. prices 278% of the 33-comparator-country average overall; 422% for brand-name drugs; U.S. generic prices below the comparator average.

Major assumptions: Ex-manufacturer (list) prices as the primary metric, with a separate rebate-adjusted brand-drug estimate; volume-weighting choices affect the precise multiplier.

Strengths: Federally commissioned, methodologically transparent, third update of a consistent long-running RAND series allowing trend comparison; not funded by an interested party on either side of the debate.

Limitations: List-price-based multipliers can overstate net U.S. spending because manufacturer rebates (substantial and non-transparent, largely captured by PBMs rather than passed to patients) are only partially reflected; does not resolve where rebate savings actually go.

Conflicts of interest: None identifiable; ASPE commissioned the study for policy planning purposes under both Democratic and Republican administrations across its several iterations.

Evidentiary support for conclusions: Strong — this is one of the most rigorously and repeatedly replicated findings in health-policy economics.

Use in policy debate: Cited by advocates of drug-price reform across the political spectrum, including in support of both the Inflation Reduction Act's Medicare negotiation provisions and single-payer proposals.

Reliability assessment: High.

---

3. Papanicolas, Woskie, and Jha — "Health Care Spending in the United States and Other High-Income Countries"

Citation: Papanicolas I, Woskie LR, Jha AK. JAMA 2018;319(10):1024–1039. DOI 10.1001/jama.2018.1150, PMID 29536101.

Year: 2018. Institution: Harvard T.H. Chan School of Public Health (published in JAMA).

Type: Cross-national comparative analysis, peer-reviewed.

Research question: What structural factors explain the gap between U.S. and peer-country health spending?

Data sources: OECD and Commonwealth Fund data, 11 high-income countries, 2013–2016.

Methodology: Comparison of spending, structural capacity, utilization, and prices across countries, decomposing the spending gap into component drivers rather than assuming higher spending equals higher volume.

Principal findings: Prices (labor, goods, pharmaceuticals) and administrative complexity, not utilization, are the primary drivers of the U.S.-peer spending gap; U.S. utilization is often lower than peer averages; administrative costs are 8% of U.S. NHE versus ~3% average elsewhere.

Major assumptions: Treats the 11-country comparator set as representative of "high-income countries" generally; relies on OECD harmonization methods that involve some cross-national definitional adjustment.

Strengths: Widely regarded as the most methodologically careful and comprehensive modern update of the "prices, not volume" literature; multiple independent data sources triangulated.

Limitations: A single cross-sectional snapshot (2013–2016); cannot establish causal direction between financing-system features and price levels; the ~3% peer-country administrative-cost average is a composite that masks meaningful cross-peer-country variation.

Conflicts of interest: None identifiable; authors are academic health-policy researchers without an institutional stake in a particular reform outcome, though Commonwealth Fund's broader institutional mission favors international benchmarking generally read as supportive of universal-coverage reform.

Evidentiary support for conclusions: Strong; considered a landmark study, frequently cited by CBO, GAO, and think tanks across the spectrum.

Use in policy debate: The single most commonly cited "why does the U.S. spend so much more" reference across the political spectrum, invoked by both single-payer advocates and market-oriented reformers (the latter to argue for price transparency and antitrust remedies rather than financing-model change).

Reliability assessment: High.

---

4. Himmelstein, Campbell, and Woolhandler — "Health Care Administrative Costs in the United States and Canada, 2017"

Citation: Himmelstein DU, Campbell T, Woolhandler S. Annals of Internal Medicine 2020;172(2):134–142. DOI 10.7326/M19-2818.

Year: 2020. Institution: CUNY Hunter College / Harvard Medical School (authors); published in Annals of Internal Medicine.

Type: Binational comparative accounting study, peer-reviewed.

Research question: How much do the U.S. and Canada spend on health care administration, and why does the gap exist?

Data sources: U.S. National Health Expenditure Accounts, Medicare cost reports for 5,526 hospitals, Canadian Institute for Health Information data, binational physician surveys, census occupational classifications.

Methodology: Bottom-up sector-by-sector accounting (insurance overhead, hospital administration, nursing home/home care, physician billing) cross-validated against workforce classification data to avoid accounting-definition distortions.

Principal findings: U.S. administrative spending $812 billion in 2017 (34.2% of measured NHE) versus 17.0% in Canada; per-capita gap exceeding $1,900; U.S. has 129.7 administrative personnel per 10,000 population versus Canada's 88.9.

Major assumptions: Assumes Canadian administrative-cost levels are an appropriate and achievable benchmark for a hypothetical U.S. single-payer system — a contestable assumption given differing labor markets, provider-payment structures, and state/provincial complexity, addressed critically in Section 5.

Strengths: Multiple independent, cross-validated data sources; consistent methodology across three decades of the authors' research program allowing trend analysis; the qualitative finding (large administrative gap) is independently corroborated by Papanicolas et al.'s very different methodology.

Limitations: The authors are co-founders of Physicians for a National Health Program, an organization advocating for single-payer reform — a direct institutional stake in producing evidence favorable to that outcome; the specific 34.2%-versus-17.0% multiplier is larger than independent estimates (Papanicolas et al.'s 8%-versus-3%), a discrepancy attributable to differing measurement scope (this study measures a broader set of administrative activities) that the paper itself explains but that critics have also used to question whether the broader scope inflates the U.S.-Canada gap; a companion critical NEJM editorial specifically challenged an earlier iteration of this research program's methodology.

Conflicts of interest: Explicit and disclosed — advocacy affiliation with PNHP.

Evidentiary support for conclusions: The directional/qualitative finding (the U.S. has substantially higher administrative costs) is well-supported and independently corroborated; the specific magnitude is contested and should be treated as an upper-bound estimate within a plausible range bounded below by Papanicolas et al.'s smaller but directionally identical estimate.

Use in policy debate: The most frequently cited figure ("$812 billion a year in paperwork") in single-payer advocacy messaging; frequently challenged by opponents on methodological grounds.

Reliability assessment: Moderate-high — directionally reliable, magnitude should be triangulated against independent sources.

---

5. Anderson, Reinhardt, Hussey, and Petrosyan / Anderson, Hussey, and Petrosyan — "It's the Prices, Stupid" and "It's Still the Prices, Stupid"

Citation: Health Affairs 22(3), 2003:89–105; Health Affairs 38(1), 2019:87–95 (DOI 10.1377/hlthaff.2018.05144, PMID 30615520).

Year: 2003, updated 2019. Institution: Johns Hopkins Bloomberg School of Public Health; Princeton (Reinhardt).

Type: Cross-national comparative analysis, peer-reviewed.

Research question: Why does the U.S. spend so much more than other OECD countries?

Data sources: OECD comparative spending, price, and utilization data.

Methodology: Direct comparison of per-capita spending against utilization measures (physicians, beds, admissions per capita) to test whether higher spending reflects higher volume.

Principal findings: The U.S. is at or below the OECD median on most utilization measures while spending far more, in both 2000 and 2016 data — the gap is attributable to prices, not volume; the 2019 update finds this pattern persisted and, if anything, intensified after the ACA.

Major assumptions: Uses aggregate utilization measures that may mask within-category intensity differences (e.g., total hospital admissions doesn't capture procedure complexity).

Strengths: The foundational study in this literature, replicated 15 years later by the same core team with the same qualitative finding, then independently replicated again by Papanicolas et al. (2018) using a different data source and research team.

Limitations: Aggregate utilization measures are a somewhat blunt instrument; a fuller decomposition (as later provided by Papanicolas et al.) offers more granularity.

Conflicts of interest: None identifiable.

Evidentiary support for conclusions: Very strong, given independent replication across three separate research efforts spanning 15 years and two research teams.

Use in policy debate: A foundational reference across the political spectrum; frequently invoked by single-payer advocates as evidence for administered pricing and by market-oriented reformers as evidence for price transparency/antitrust remedies instead.

Reliability assessment: High.

6. Congressional Budget Office — "Economic Effects of Five Illustrative Single-Payer Health Care Systems"

Citation: CBO Working Paper 2022-02 (Jaeger Nelson), Feb. 2022. https://www.cbo.gov/publication/57637

Year: 2022. Institution: U.S. Congressional Budget Office (nonpartisan legislative-branch agency).

Type: Microsimulation/general-equilibrium economic model.

Research question: What would be the economic and budgetary effects of illustrative single-payer designs?

Data sources: CBO's health insurance simulation model, integrated with a general-equilibrium overlapping-generations macroeconomic model.

Methodology: Five illustrative designs varying provider payment rates, cost-sharing levels, and long-term-care coverage, modeled to a 2030 endpoint assuming hypothetical 2025 implementation.

Principal findings: Administrative savings of $533–562 billion (≈1.8% of GDP) in 2030 across all designs; national health expenditure effects ranging from a $743 billion decrease to a $290 billion increase depending on the payment-rate and cost-sharing assumptions; federal subsidies increasing by $1.5–3.0 trillion.

Major assumptions: Static federal-financing assumptions (does not specify a particular tax package); induced-demand elasticities drawn from prior health-economics literature (including the Oregon and RAND health insurance experiments); assumes providers behave as modeled rather than exiting practice or reducing hours in response to lower payment rates.

Strengths: The single most methodologically transparent, most frequently updated, and most institutionally neutral quantitative model in the literature; explicitly designed to show how sensitive results are to specific assumptions rather than to produce one "answer."

Limitations: "Illustrative" by design — does not score any actual pending legislation; a dynamic macro model is only as good as its behavioral-response assumptions, which cannot be directly validated because no country has implemented a system directly comparable to the U.S. economy's scale and structure.

Conflicts of interest: None; CBO's institutional design (a nonpartisan agency serving both parties in Congress) is specifically intended to prevent this.

Evidentiary support for conclusions: The paper's central conclusion — that the sign and magnitude of the total spending effect depends almost entirely on provider payment rates and induced demand, not on administrative savings (which are comparatively stable across scenarios) — is well-supported by its own sensitivity analysis and independently corroborated by the pattern seen across the other cost models reviewed here (RAND, Urban Institute, Mercatus, PERI).

Use in policy debate: The most frequently cited neutral reference point by journalists, other government agencies, and researchers across the spectrum when discussing single-payer cost ranges.

Reliability assessment: Very high, understood correctly as an assumption-sensitivity map rather than a single forecast.

---

7. RAND Corporation — "An Assessment of the New York Health Act"

Citation: Liu JL, White C, Nowak SA, Wilks A, Ryan JL, Eibner C. RAND Research Report RR-2424-NYSHF (2018). https://www.rand.org/pubs/research_reports/RR2424.html

Year: 2018. Institution: RAND Corporation, commissioned by the New York State Health Foundation.

Type: State-level microsimulation.

Research question: What would be the fiscal and coverage effects of a state-level single-payer plan in New York?

Data sources: State claims and demographic data, calibrated microsimulation model.

Methodology: Projected total health spending and required state tax revenue for 2022, 2026, and 2031 under the proposed New York Health Act design.

Principal findings: Total health spending roughly flat in 2022, approximately 3% lower by 2031 (cumulative ~2% ten-year savings), contingent on achieving assumed administrative savings and provider-payment restraint; would require an additional $139 billion in state tax revenue in 2022 (a 156% increase over projected state tax revenue), and depends on federal waivers being granted to fold Medicare and Medicaid funding into the state plan.

Major assumptions: Assumes federal government approval of waivers allowing New York to redirect federal Medicare/Medicaid dollars into the state plan — an assumption RAND's own authors characterized as uncertain, and one later commentators (e.g., the Empire Center) argued was politically "iffy," particularly under an unsympathetic federal administration.

Strengths: Rigorous, transparent microsimulation from a nonpartisan research organization with no stake in the outcome; explicit about the waiver-dependency risk.

Limitations: State-level analysis cannot be directly generalized to a national system, since a national system would not face the same interstate risk-pool, ERISA-preemption, or federal-waiver dependency issues; a single state's estimate is also more exposed to adverse selection and outmigration risk than a national estimate would be.

Conflicts of interest: Funded by a nonpartisan state health foundation with an interest in state health policy generally, not tied to advocacy for or against single payer specifically; RAND is a long-standing nonpartisan research organization.

Evidentiary support for conclusions: The finding that fiscal outcomes are highly contingent on politically uncertain federal cooperation is well-supported by the study's own sensitivity analysis.

Use in policy debate: Widely cited in the New York state legislative debate over the NY Health Act, and as a data point in the broader national feasibility debate over state-level single-payer implementation.

Reliability assessment: High for what it models (a specific state proposal under specific assumptions); of limited direct relevance to national-level feasibility questions.

---

8. Mercatus Center — "The Costs of a National Single-Payer Healthcare System"

Citation: Blahous C. Mercatus Center Working Paper, George Mason University, July 30, 2018.

Year: 2018. Institution: Mercatus Center, a free-market-oriented research center at George Mason University substantially associated with Koch-network funding (reported by multiple outside sources including The Intercept; Mercatus's own working-paper page does not itemize funding).

Type: Actuarial/budgetary cost model.

Research question: What would the Medicare for All Act of 2017 (S.1804) cost the federal government?

Data sources: CMS actuarial projections, adapted with the author's own assumptions about administrative savings, drug-price reductions, and provider payment rates.

Methodology: Ten-year (2022–2031) federal budget cost projection, explicitly incorporating optimistic assumptions favorable to the bill (drawn from Medicare for All supporters' own claims) as an upper bound of favorable assumptions, then testing sensitivity.

Principal findings: Total federal spending increases by approximately $32.6 trillion over 2022–2031, even incorporating supporters' own optimistic administrative-savings and drug-price assumptions; the author explicitly frames this as a conservative floor, arguing actual costs would likely be higher because the assumed 40%-plus cut to provider payment rates is probably economically and politically unsustainable.

Major assumptions: A greater-than-40% reduction in provider payment rates relative to current private-insurance rates, which the author himself flags as an unrealistic assumption baked in specifically because it was advocates' own preferred assumption, not because Blahous considers it plausible; this dual role (both advocate's-best-case input and the study's own conclusion that even the best case looks expensive) is central to correctly interpreting the paper.

Strengths: Transparent about its assumptions and their sourcing; the author has subsequently and publicly corrected misuse of his own findings (Manhattan Institute, 2019), specifically rebutting the claim, spread by some outlets and political figures, that his study found aggregate national health spending would fall — clarifying that his estimate concerns the *federal budget*, not total national health spending, and that this distinction was frequently lost in political commentary on both sides.

Limitations: Explicit institutional and ideological orientation (free-market, Koch-network-associated) toward opposing an expanded government role in health financing; the deliberate choice to use advocates' own optimistic assumptions as an upper bound, while methodologically defensible as a "best case for supporters" framing, produces a headline number that is easily quoted out of that context (as happened repeatedly in press coverage).

Conflicts of interest: Explicit and disclosed — institutional funding and orientation opposed to expanded government health financing.

Evidentiary support for conclusions: The federal-cost estimate itself is broadly corroborated by the independently produced, institutionally opposite-leaning Urban Institute estimate (~$32 trillion, 2016) — a striking convergence given the two institutions' opposing orientations, and one of the more evidentially solid figures in this entire review precisely because it did not depend on institutional agreement.

Use in policy debate: The most frequently cited "cost of Medicare for All" figure in opposition messaging; frequently mischaracterized (per Blahous's own correction) as evidence of a net *national* spending increase, when the study's aggregate national spending estimate cannot be cleanly separated from the federal-budget estimate without additional analysis Blahous did not fully provide in the original paper.

Reliability assessment: Moderate-high for the federal-budget estimate specifically (corroborated by Urban Institute); the paper's rhetorical framing invites overreading beyond what the numbers themselves establish.

---

9. Urban Institute — "The Sanders Single-Payer Health Care Plan: The Effect on National Health Expenditures and Federal and Private Spending"

Citation: Holahan J, Clemans-Cope L, Buettgens M, Favreault M, Blumberg LJ, Ndwandwe S. Urban Institute Health Policy Center, May 2016.

Year: 2016. Institution: Urban Institute, a Washington think tank generally regarded as center-left/institutionally independent, funded in this instance in part by the Commonwealth Fund and RWJF among other funders.

Type: Microsimulation cost model.

Research question: What would the 2016 Sanders single-payer plan cost, and how would costs be distributed?

Data sources: Urban's Health Insurance Policy Simulation Model (HIPSM), calibrated to CPS and MEPS survey data.

Methodology: Ten-year (2017–2026) projection of national and federal spending under a specific, comprehensive benefit design (zero cost-sharing, inclusion of long-term services and supports, dental, vision, hearing).

Principal findings: Federal spending increases $32.0 trillion (232.7%) over ten years; national health expenditure increases $6.6 trillion (16.6%) over ten years.

Major assumptions: Zero patient cost-sharing (a major driver of both the induced-demand and federal-cost estimates); administrative costs fixed at 6% of claims (i.e., assumes only a modest reduction, not the deeper cuts PNHP-affiliated researchers argue are achievable); assumes providers accept payment near Medicare rates for physicians and near-full cost reimbursement for hospitals.

Strengths: A rigorously documented, widely used, and frequently updated microsimulation model with a long track record in health-policy analysis (also used by federal agencies and other researchers); assumptions are explicit and separately testable.

Limitations: Assumed zero net provider-side administrative savings, a choice PNHP-affiliated critics (Himmelstein and Woolhandler) argue is unrealistic and understates savings by roughly $250 billion; assumed a 20.6% utilization increase that the same critics argue is implausible relative to the historical Medicare rollout precedent (1965–66), which saw no comparable jump in physician visits.

Conflicts of interest: None indicating advocacy for or against single payer specifically; Urban Institute is not aligned with either side of this debate, though its general institutional orientation favors evidence-based incremental reform, and it has been publicly and specifically criticized by PNHP for what PNHP characterizes as systematically conservative administrative-savings assumptions.

Evidentiary support for conclusions: The federal-cost estimate is well-corroborated by the independently produced, oppositely-oriented Mercatus estimate; the administrative-savings assumption is the most contested single input and is addressed directly (with opposing views) in Section 5.

Use in policy debate: Widely cited by both single-payer opponents (for the large federal-cost figure) and, more selectively, by supporters of more incremental reform as evidence that comprehensive single-payer requires a very large federal tax increase even under a moderate administrative-savings assumption.

Reliability assessment: High for the modeling itself; the size of the administrative-savings assumption is a genuinely open and contested empirical question, addressed critically in Section 5.

---

10. Cai, Runte, Kahn, and colleagues — "Projected Costs of Single-Payer Healthcare Financing in the United States: A Systematic Review of Economic Analyses"

Citation: Cai C, Runte J, Ostrer I, Berry K, Ponce N, Rodriguez M, Bertozzi S, White JS, Kahn JG. PLOS Medicine 2020;17(1):e1003013. DOI 10.1371/journal.pmed.1003013.

Year: 2020. Institution: UCSF, UCLA, and UC Berkeley (authors); published in PLOS Medicine.

Type: Systematic review of economic analyses (a review of the literature itself, not a new primary cost model).

Research question: What do the published quantitative single-payer cost analyses collectively conclude?

Data sources: 22 credible economic models drawn from 18 studies of single-payer proposals (8 national, 14 state-level), published 1991–2018.

Methodology: Systematic literature search and structured synthesis (not a meta-analysis with pooled effect sizes, given the heterogeneity of the underlying models).

Principal findings: 86% (19 of 22) of the reviewed models projected net savings in the first year of implementation, with a median first-year savings of 3.46% of total health-care costs; 91% projected savings would grow over time; the single largest identified savings driver across models was administrative simplification (median estimated savings of 8.8%).

Major assumptions: Inherits the assumptions of each underlying model reviewed; does not independently adjudicate which underlying assumptions are more realistic, meaning the "86% show savings" headline finding reflects the composition of the studies that have been conducted and published, not an independent assessment of which are correct.

Strengths: Rigorous, pre-specified systematic review methodology; useful for showing that net savings are a common, not fringe, conclusion across the diverse published literature.

Limitations: A systematic review of a body of studies that itself skews toward certain assumption sets cannot resolve disagreements about which assumptions are more realistic; publication and selection effects in what gets modeled and published are not addressed; the review received financial support from PNHP, and the lead author held a leadership role in Students for a National Health Program, a direct institutional tie to single-payer advocacy that, while disclosed, is directly relevant given that a systematic review's conclusions depend heavily on which underlying studies are characterized as "credible" for inclusion.

Conflicts of interest: Disclosed PNHP funding and leadership tie for the lead author.

Evidentiary support for conclusions: The finding that most published models show savings is accurately reported, but "most published models agree" is a different and weaker claim than "the true effect is savings" — a distinction obscured when this review is cited (as it frequently is) as though it independently proves single payer would save money, rather than showing that most modelers who chose to model single payer chose assumption sets that produce savings.

Use in policy debate: Frequently cited by single-payer advocates as "22 studies agree" evidence.

Reliability assessment: Moderate — reliable as a description of the published literature's central tendency, but should not be read as independent confirmatory evidence beyond what the underlying (heterogeneous, assumption-dependent) studies themselves establish.

---

11. Galvani, Parpia, Foster, Singer, and Fitzpatrick — "Improving the Prognosis of Health Care in the USA"

Citation: Lancet 2020;395(10225):524–533. DOI 10.1016/S0140-6736(19)33019-3.

Year: 2020. Institution: Yale School of Public Health, Center for Infectious Disease Modeling and Analysis.

Type: Epidemiological-economic model, peer-reviewed.

Research question: What would be the cost and mortality effects of the Medicare for All Act of 2017?

Data sources: CMS spending data, published mortality and access-to-care literature.

Methodology: Modeled cost savings (administrative, payment-rate, and primary-care-access channels) and translated estimated access improvements into projected mortality reductions using external mortality-effect estimates from the literature.

Principal findings: 13% reduction in national health expenditure (~$450 billion/year); approximately 68,000 fewer deaths and 1.73 million additional life-years per year.

Major assumptions: Hospitals accept Medicare payment rates as a system-wide ceiling despite currently often losing money on Medicare patients; zero increase in utilization among the already-insured despite the complete elimination of their cost-sharing; the 68,000-lives estimate rests on a single external mortality-effect study rather than a range of estimates from the broader literature.

Strengths: Peer-reviewed in a top general-medicine journal; attempts a genuinely comprehensive integration of cost and mortality modeling in one framework, which few other studies do.

Limitations: Independent, non-partisan fact-checking organizations (PolitiFact) and journalists (Reason) identified the zero-utilization-increase-among-already-insured assumption as inconsistent with standard health-economics findings (insurance economics predicts a positive utilization response to eliminated cost-sharing, sometimes called the "moral hazard" effect, documented in the RAND Health Insurance Experiment and elsewhere); the reliance on a single external mortality study rather than a range is a further limitation; the model excludes roughly $4 trillion in long-term-care costs that the underlying Sanders legislation would actually have covered, which — if included — would substantially change the net cost conclusion.

Conflicts of interest: None explicitly disclosed as institutional advocacy, though the study's framing and headline results have been prominently used in single-payer political advocacy (including by the Sanders campaign), and its methodology choices consistently favor the more favorable end of plausible assumptions.

Evidentiary support for conclusions: The magnitude of both the cost-savings and mortality claims substantially exceeds what more methodologically conservative models (CBO, Urban Institute) find, and independent fact-checkers have specifically identified assumption choices that inflate the headline results; this study should be treated as an upper-bound, optimistic-case estimate rather than a central estimate.

Use in policy debate: Widely cited by single-payer advocates, including in the 2020 Democratic presidential primary, as evidence Medicare for All would both save money and save lives; also widely cited by critics as an example of an outlier, assumption-driven result.

Reliability assessment: Low-moderate; directionally consistent with other studies on some channels (administrative savings, payment-rate effects) but the specific magnitude, especially the mortality claim, is not well-supported once the identified assumption problems are accounted for.

---

12. Cooper, Craig, Gaynor, and Van Reenen — "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured"

Citation: NBER Working Paper 21815 (2015, rev. 2018); Quarterly Journal of Economics 134(1), 2019:51–107. PMID 32981974.

Year: 2019 (published version). Institution: Yale, University of Pennsylvania (Wharton), Carnegie Mellon, MIT/LSE.

Type: Empirical claims-data analysis, peer-reviewed, quasi-experimental (merger analysis).

Research question: What explains the large geographic variation in privately-insured health spending, and what role does hospital market concentration play?

Data sources: Health Care Cost Institute claims data covering 28% of U.S. employer-sponsored-insurance enrollees; data on 366 hospital mergers, 2007–2011.

Methodology: Decomposition of geographic spending variation into price versus quantity components; difference-in-differences analysis of price changes following hospital mergers, using merger distance as an instrument for market-power change.

Principal findings: Privately insured per-beneficiary spending varies threefold across U.S. markets, weakly correlated with Medicare spending in the same areas; roughly half of this variation is price-driven; monopoly hospitals charge 12–15% more than hospitals in markets with four or more competing systems; mergers of hospitals within five miles of each other raised prices more than 6%, while mergers of hospitals more than 25 miles apart had no significant price effect.

Major assumptions: Uses merger distance as a proxy for the degree of competitive overlap, a standard but not unassailable identification strategy in industrial-organization economics.

Strengths: Large, high-quality administrative claims dataset; credible quasi-experimental identification strategy (distance-based merger analysis) rather than simple cross-sectional correlation; published in a top economics journal.

Limitations: Covers privately insured populations only, not Medicare or Medicaid; cannot speak directly to a single-payer counterfactual, only to the role of market power within the current multipayer system.

Conflicts of interest: None identifiable.

Evidentiary support for conclusions: Strong, given the quasi-experimental design.

Use in policy debate: Cited by both antitrust-oriented reformers (as evidence for stronger merger enforcement within the current system) and single-payer advocates (as evidence that administered/negotiated pricing, which single payer would impose, addresses a real and quantifiable market-power problem).

Reliability assessment: High.

---

13. Mello, Chandra, Gawande, and Studdert — "National Costs of the Medical Liability System"

Citation: Health Affairs 29(9), 2010:1569–1577. DOI 10.1377/hlthaff.2009.0807, PMID 20820010.

Year: 2010. Institution: Harvard, Stanford (Chandra), Harvard Medical School (Gawande), University of Melbourne/Stanford (Studdert).

Type: National cost accounting/estimation study, peer-reviewed.

Research question: What is the total cost of the U.S. medical liability system, including defensive medicine?

Data sources: Malpractice indemnity payment data, insurer administrative-cost data, and prior defensive-medicine cost literature.

Methodology: Aggregation of indemnity payments, legal/administrative defense costs, and estimated defensive-medicine spending into a national total.

Principal findings: Total medical liability system cost of $55.6 billion per year (2008 dollars), 2.4% of total U.S. health spending; defensive medicine specifically estimated at $45.6 billion per year. The authors' own conclusion, notable for its lack of advocacy framing in either direction, is that malpractice/defensive-medicine reform, while "not trivial," is "unlikely to be a source of significant savings" at the system level.

Major assumptions: Defensive-medicine estimation necessarily requires inferring physician intent (ordering a test partly or wholly to avoid liability risk rather than for clinical benefit), which is inherently difficult to measure precisely and relies on physician self-report and literature-based approximation.

Strengths: A rare example of a rigorously quantified estimate of a frequently invoked but rarely quantified cost driver; the authors' own conclusion cuts against the popular narrative (on the political right) that malpractice reform is a major lever for cost control, lending credibility to the finding as not being shaped by a predetermined narrative.

Limitations: Defensive-medicine estimation is inherently imprecise; the study is now over 15 years old and health-care prices/practice patterns have changed since, though no more recent comparably rigorous national estimate was identified in this review.

Conflicts of interest: None identifiable.

Evidentiary support for conclusions: Strong, and notably self-limiting (the authors do not claim their finding supports a larger policy conclusion than the data warrant).

Use in policy debate: Frequently cited to push back against claims, common in the tort-reform debate, that malpractice costs are a primary driver of high U.S. health spending; directly relevant to this review's finding that the international spending gap is overwhelmingly a story about pharmaceutical, hospital, and administrative prices rather than about liability costs.

Reliability assessment: High.

---

14. Finkelstein, Baicker, and colleagues — The Oregon Health Insurance Experiment

Citation: Finkelstein A, Taubman S, Wright B, et al. "The Oregon Health Insurance Experiment: Evidence from the First Year." Quarterly Journal of Economics 2012;127(3):1057–1106; Baicker K, Taubman SL, Allen HL, et al. "The Oregon Experiment — Effects of Medicaid on Clinical Outcomes." NEJM 2013;368:1713–1722.

Year: 2012–2013. Institution: MIT (Finkelstein), Harvard (Baicker).

Type: Randomized controlled trial (a genuine lottery-based natural experiment) — the gold standard of causal-inference evidence in this entire review.

Research question: What is the causal effect of Medicaid coverage on health care utilization, financial strain, and health outcomes?

Data sources: Oregon's 2008 Medicaid expansion lottery, which randomly selected a limited number of applicants from a waitlist, creating treatment and control groups by random assignment.

Methodology: Intent-to-treat and treatment-effect analysis comparing lottery winners to lottery losers roughly two years post-randomization.

Principal findings: Lottery winners showed increased health care utilization, essentially eliminated catastrophic medical expenditures, and reduced rates of depression; no statistically significant effect was detected on several physical-health biomarkers (blood pressure, cholesterol, glycated hemoglobin) within the roughly two-year study window.

Major assumptions: None beyond standard RCT assumptions (which held well given the genuine randomization); the null biomarker findings are frequently over-interpreted.

Strengths: True randomization — the strongest possible causal-inference design available in health-insurance research, and one that essentially cannot be replicated at national scale (no country would randomize coverage nationally), making this a uniquely valuable natural experiment.

Limitations: The study is statistically underpowered to detect the biomarker and mortality effects it examined, given the size of the lottery sample and the relatively short follow-up window — the correct statistical interpretation of the null biomarker results is "not enough evidence to detect an effect of the size expected," not "insurance has no effect on physical health," a distinction frequently and consequentially lost in political debate on both sides. The study is also specific to Oregon's Medicaid population and benefit design in 2008, which may not generalize to a national universal-coverage program with a different benefit structure or population.

Conflicts of interest: None identifiable.

Evidentiary support for conclusions: Very strong for what the study actually found (utilization, financial protection, and mental-health effects); appropriately weak for the claim, sometimes advanced by insurance skeptics, that "the study proves insurance doesn't improve physical health" — that claim is not supported by the study's own power analysis.

Use in policy debate: Cited by both sides — by single-payer skeptics for the null biomarker finding (often without the power caveat) and by insurance-expansion advocates for the financial-protection and mental-health findings.

Reliability assessment: Very high for the effects it was powered to detect; frequently misused as evidence for effects it was not powered to detect.

---

15. National Research Council and Institute of Medicine — "U.S. Health in International Perspective: Shorter Lives, Poorer Health"

Citation: Woolf SH, Aron L, eds. National Academies Press, 2013. PMID 24006554.

Year: 2013. Institution: National Academies of Sciences, Engineering, and Medicine (a congressionally chartered, nonpartisan scientific advisory body).

Type: Expert consensus panel report, synthesizing existing literature (not new primary research).

Research question: Why do Americans live shorter, less healthy lives than people in other high-income countries?

Data sources: Comprehensive review of the epidemiological, health-services, and social-determinants literature comparing the U.S. to 16 peer countries.

Methodology: Multidisciplinary expert panel synthesis, examining behavioral factors (obesity, historical smoking, drug and alcohol use, firearm violence), social factors (poverty, inequality, social mobility), environmental factors, and health-system factors.

Principal findings: The U.S. health disadvantage is real, longstanding, and pervasive across age groups, but is not attributable to any single cause; the panel explicitly declines to offer a single clean percentage-based attribution across categories, emphasizing instead that these factors interact.

Major assumptions: None beyond standard evidence-synthesis practice; the panel's explicit refusal to force a false precision onto an inherently multi-causal problem is itself a methodological strength worth noting.

Strengths: The single most authoritative, comprehensive, and methodologically cautious treatment of the "why is the U.S. sicker" question, produced by a body specifically chartered to be nonpartisan and evidence-driven.

Limitations: As a synthesis rather than new primary research, its conclusions are only as strong as the underlying literature it reviews, much of which is itself correlational; more than a decade old, though a fully comparable, equally authoritative update was not identified in this review.

Conflicts of interest: None identifiable; National Academies committee members undergo formal conflict-of-interest vetting.

Evidentiary support for conclusions: Very strong for the qualitative conclusion (multi-causal, not primarily a health-care-system-quality failure); appropriately cautious about assigning precise percentage weights.

Use in policy debate: Cited across the spectrum, sometimes selectively — single-payer skeptics cite it to argue the U.S. outcomes gap isn't primarily a financing problem, while access-focused researchers cite the report's health-system-factors chapter (unmet need, financial barriers to care) as evidence that financing does matter for the subset of the gap attributable to access.

Reliability assessment: Very high.

---

16. Commonwealth Fund — "Mirror, Mirror 2024: A Portrait of the Failing U.S. Health System"

Citation: Blumenthal D, et al. Commonwealth Fund, 2024. DOI 10.26099/ta0g-zp66.

Year: 2024. Institution: The Commonwealth Fund, a private nonpartisan foundation with a long-standing institutional mission of promoting a high-performing, equitable health system, frequently (and explicitly) oriented toward international benchmarking that has consistently found the U.S. system underperforming since the series began in 2004.

Type: Composite cross-national ranking/scorecard.

Research question: How does U.S. health system performance compare to nine other high-income countries across access, care process, administrative efficiency, equity, and health outcomes?

Data sources: OECD, WHO, and Commonwealth Fund's own International Health Policy Survey data.

Methodology: Composite scoring across five domains using a weighted set of indicators.

Principal findings: The U.S. ranks last overall among the ten countries studied, last specifically on outcomes and equity, but second on care-process measures.

Major assumptions: Composite scoring necessarily involves subjective weighting choices about which indicators matter and how much; different weighting schemes could plausibly produce different rankings, though the U.S.'s poor performance on outcomes and equity specifically (not just the composite) is robust across the underlying component data.

Strengths: Long-running, methodologically consistent series allowing trend comparison since 2004; transparent about its component indicators, allowing readers to examine the underlying data rather than relying solely on the composite rank.

Limitations: Commonwealth Fund's institutional mission (explicitly oriented toward advancing a "high-performing, equitable health system," historically understood as favoring more universal-coverage-oriented reform) means its choice of framing ("failing U.S. health system") and indicator selection should be read with that orientation in mind, even though the underlying component data (which are separately verifiable against OECD/WHO sources) support the qualitative findings on outcomes and equity specifically; the composite "last

place" ranking is more a function of weighting choices than any single indicator, and the report itself shows the U.S. performing well on some component measures (care process).

Conflicts of interest: Institutional mission orientation, disclosed by the Fund itself; not a conflict in the financial sense but a relevant framing consideration.

Evidentiary support for conclusions: The individual component findings (poor performance on access, equity, and amenable mortality specifically) are well-supported by the underlying data and independently corroborated elsewhere in this review; the composite "worst overall" headline is more contestable and sensitive to methodology.

Use in policy debate: The single most frequently cited "the U.S. ranks last" reference in single-payer and universal-coverage advocacy.

Reliability assessment: High for the component findings; moderate for the composite ranking, which should be read as "the U.S. underperforms substantially on access, equity, and outcomes" rather than "the U.S. health system is worst in every respect."

---

17. The Heritage Foundation — "How Medicare for All Harms Working Americans"

Citation: Haislmaier E, Hall JB. The Heritage Foundation, Nov. 19, 2019.

Year: 2019. Institution: The Heritage Foundation, a self-described conservative think tank with historically undisclosed current donors (2023 revenue ~$101 million).

Type: Static distributional accounting analysis.

Research question: How would a Medicare for All-style payroll-tax financing scheme affect household finances?

Data sources: 2016 Medical Expenditure Panel Survey (MEPS) data.

Methodology: Static accounting — applies a hypothetical payroll-tax financing scheme to existing household spending patterns without modeling behavioral responses (e.g., wage effects of employer payroll-tax incidence, changed consumption patterns, or macroeconomic feedback).

Principal findings: 73.5% of Americans would be financially worse off under the modeled financing scheme; average household disposable income down $5,671 per year (11%); federal spending up $2.387 trillion in 2020 (a 50% increase).

Major assumptions: A specific, single financing mechanism (a particular payroll tax structure) is assumed rather than tested against alternatives; critically, a purely static framework does not model employer behavior (e.g., whether employers would raise wages if relieved of premium-payment obligations, a standard assumption in the broader labor-economics literature on employer-sponsored insurance incidence) — a limitation that, if incorporated, would likely change the household-income-effect estimate substantially in the reform-favorable direction, and one the report does not address.

Strengths: Transparent about its specific financing-mechanism assumption and data source.

Limitations: Static rather than dynamic modeling is a significant limitation for a scenario (large tax-and-spending shift) where behavioral and general-equilibrium responses are large relative to the direct effect being measured; does not model alternative financing mixes (progressive income taxation, wealth taxes, financial-transaction taxes) that many actual Medicare for All proposals include alongside or instead of payroll taxes.

Conflicts of interest: Explicit institutional and ideological orientation opposed to expanded government health financing.

Evidentiary support for conclusions: The finding is highly sensitive to the single financing mechanism chosen and the static (non-behavioral) modeling approach; a different, equally plausible financing mix or a dynamic model incorporating wage pass-through could produce materially different distributional conclusions.

Use in policy debate: Frequently cited in conservative opposition messaging to Medicare for All.

Reliability assessment: Low-moderate; the arithmetic is not in question, but the framing (single financing assumption, no behavioral modeling, no consideration of alternative financing mixes) limits how much weight the finding can bear as a general statement about "Medicare for All."

---

18. Committee for a Responsible Federal Budget — "Choices for Financing Medicare for All"

Citation: Committee for a Responsible Federal Budget, March 17, 2020.

Year: 2020. Institution: CRFB, a self-described nonpartisan fiscal watchdog organization with a documented funding relationship with the Peter G. Peterson Foundation (a foundation with a long-standing institutional focus on deficit reduction).

Type: Financing-options budgetary analysis, incorporating third-party dynamic macroeconomic modeling (Penn Wharton Budget Model).

Research question: What tax or financing changes would be required to fund Medicare for All, and what would their economic effects be?

Data sources: CBO and other cost estimates (including the sources reviewed above), Penn Wharton Budget Model simulations.

Methodology: Presents multiple illustrative financing packages (payroll tax, income surtax, VAT, per-capita premium, corporate/income tax increases, spending cuts, deficit financing) each sized to close an assumed $25–35 trillion ten-year financing gap, then reports each package's macroeconomic effects using Penn Wharton's model.

Principal findings: A $25–35 trillion (midpoint ~$30 trillion) ten-year financing need; payroll-tax financing would reduce GDP by 7.3% by 2030 (Penn Wharton estimate) versus 5.9% for deficit financing and 2.3% for premium financing; most tax-financing options are more progressive than current law, while a flat per-capita premium would be regressive.

Major assumptions: Adopts a specific ten-year cost range drawn from the broader literature (including Mercatus and Urban Institute) rather than producing an independent primary cost estimate; the Penn Wharton macroeconomic effects depend on that model's own structural assumptions about labor supply and capital responses to tax changes, which are contested within economics more broadly (as is true of any dynamic macro model).

Strengths: Notably even-handed in its explicit finding that most financing options are more progressive than the current system — a conclusion that cuts against a simple "single payer means higher middle-class taxes" narrative, and one CRFB reports despite its institutional deficit-focus orientation.

Limitations: Relies on other institutions' underlying cost estimates rather than an independently produced primary cost model; institutional orientation toward deficit reduction generally, which shapes how the report frames financing "choices" (as costs to be minimized) rather than, say, comparing financing burdens to the household costs of the current system it would replace.

Conflicts of interest: Documented funding relationship with a foundation whose named institutional focus is deficit reduction; CRFB's orientation is more accurately described as deficit-hawk than left/right on coverage policy specifically.

Evidentiary support for conclusions: The progressivity finding across most financing options is a genuinely evidence-based and somewhat underappreciated conclusion, given the report's institutional starting point; the headline $25–35 trillion figure is not independently derived but is a reasonable synthesis of the broader cost-estimate literature reviewed above.

Use in policy debate: Widely cited in mainstream reporting on Medicare for All financing options; criticized by PNHP as failing to compare financing costs to current-system costs on an apples-to-apples basis.

Reliability assessment: Moderate-high for the financing-options analysis; appropriately read as a synthesis of others' cost estimates rather than independent primary evidence on the cost question itself.

---

19. Government Accountability Office — "Canadian Health Insurance: Lessons for the United States" and "Estimating Costs and Savings for the United States"

Citation: GAO/HRD-91-90 (1991); GAO/HRD-92-83 (1992).

Year: 1991–1992. Institution: U.S. Government Accountability Office (nonpartisan legislative-branch agency).

Type: Government program evaluation and comparative costing analysis.

Research question: What can the U.S. learn from Canada's single-payer system, and what would adopting a similar system cost or save?

Data sources: Canadian government health-spending and administrative data, U.S. National Health Expenditure data of the period.

Methodology: Direct binational administrative-cost and coverage comparison, commissioned at congressional request.

Principal findings: Found substantial administrative-cost savings potential in a Canadian-style approach, based on early-1990s data, and provided one of the earliest rigorous nonpartisan quantifications of this question — a precursor to the more recent Himmelstein/Woolhandler and Papanicolas/Woskie/Jha research programs.

Major assumptions: Reflects early-1990s Canadian and U.S. health-system structures, which have changed materially in the intervening three decades (Canadian provincial cost-control mechanisms, U.S. managed-care penetration, and administrative technology have all evolved).

Strengths: Nonpartisan, congressionally commissioned, methodologically direct; historically important as one of the earliest serious quantitative treatments of this question by a U.S. government body.

Limitations: More than three decades old; specific dollar figures are of limited direct relevance today, though the qualitative administrative-cost-gap finding has been repeatedly reaffirmed by later research.

Conflicts of interest: None identifiable.

Evidentiary support for conclusions: The qualitative finding (Canada's system has structurally lower administrative costs) has held up well against three subsequent decades of independent research; the specific quantitative estimates from 1991–92 should not be relied upon today.

Use in policy debate: Primarily of historical interest today; superseded in practical citation frequency by the more recent Himmelstein/Woolhandler and Papanicolas et al. work, though it establishes that this question has been studied by nonpartisan government bodies for over three decades with a consistent qualitative answer.

Reliability assessment: High for its qualitative conclusion; outdated for specific figures.

Sources Explicitly Not Included Above Despite Being Sought

Two items referenced in early drafts of this review's research plan could not be verified and are explicitly not cited as fact anywhere in this report: (1) a purported GAO report compiling and comparing approximately 20 single-payer cost studies — extensive search of GAO's own publication database found no such report, and this appears to be a conflation with the Cai et al. PLOS Medicine systematic review of 22 models; (2) a RAND analysis of Colorado's 2016 Amendment 69 ballot measure — that analysis was in fact conducted by the Colorado Health Institute, a separate organization, not RAND. Both corrections are noted here in the interest of transparency about the research process and to avoid the citation of studies that do not exist as described.

5. Critical Evaluation of the Major Analyses

Rather than repeating source-by-source critique already provided in the annotated bibliography, this section applies a consistent set of analytical questions across all major cost models, to make visible exactly which assumptions are doing the work in each one.

What assumptions drive the result? Across every model reviewed — CBO, RAND, Urban Institute, Mercatus, PERI, and Galvani et al. — the single largest driver of the bottom-line national-spending result is the assumed provider payment rate (what hospitals and physicians are paid relative to current private-insurance rates), not the size of administrative savings. Administrative-savings estimates cluster in a comparatively narrow range across most models (roughly 1.5–8% of total spending, depending on scope and whether provider-side savings are included), while assumed provider payment rates range from close to current private rates (producing higher total spending) to deep cuts approaching or below Medicare rates (producing lower total spending) — a much wider relative range with much larger dollar consequences. CBO's own Working Paper 2022-02 makes this point explicitly: its five scenarios differ almost entirely along the payment-rate and cost-sharing dimensions, and the resulting national-spending effect swings by more than $1 trillion depending on which combination is chosen, while the administrative-savings component stays comparatively stable across all five.

What variables are included or excluded? A recurring and consequential inclusion/exclusion problem is long-term services and supports (LTSS). The Galvani/Yale model's favorable 13% spending-reduction finding explicitly excludes roughly $4 trillion in LTSS costs that the underlying Sanders legislation would have covered — a scope choice that, if corrected, would materially change the study's own conclusion. The Urban Institute and Mercatus estimates, by contrast, include LTSS, which is one reason their federal-cost figures are dramatically higher than Galvani's. This is not a left-right methodological dispute so much as a scope-definition dispute: studies that model a narrower benefit package show smaller cost increases (or larger savings) than studies that model the full, LTSS-inclusive legislative text, essentially by definition.

Are counterfactual assumptions reasonable? The Urban Institute's assumption of zero net provider-side administrative savings, and the Himmelstein/Woolhandler-derived assumption (used by PERI and cited in PNHP's critiques of Urban) that U.S. providers could achieve Canadian-level administrative efficiency, are both defensible in different ways and both somewhat unrealistic in different ways: Urban's assumption ignores well-documented, bottom-up evidence (the Tseng et al. micro-costing study) that U.S. provider billing overhead is real and substantial and would very plausibly fall under payment simplification; the PNHP-derived assumption that the U.S. could rapidly reach Canadian per-capita administrative-cost levels likely understates the transition friction, workforce retraining and redeployment costs, and structural differences (state-by-state licensure and

regulation, more complex U.S. hospital ownership structures, existing health IT systems built around multipayer billing) that would slow such convergence, at least in the near term. CBO's more moderate administrative-savings assumption (falling to roughly 1.5–2% of spending, well above Canada's public-plan-only 1.3% but well below the current 12% private-insurer average) is a genuinely intermediate, defensible position between these two poles, which is one reason this review treats CBO's estimates as the most useful central reference point.

How sensitive are the conclusions to changes in assumptions? Extremely sensitive, and this sensitivity is the single most important, well-documented, and underappreciated finding of this entire literature. CBO's own Working Paper 2022-02 exists specifically to demonstrate this sensitivity by construction. The fact that Mercatus and Urban Institute — institutions with opposite ideological starting points — arrived at similar federal-cost figures (~$32–33 trillion) for a comparable, comprehensive, LTSS-inclusive benefit design, while Galvani's narrower-benefit, more-optimistic-assumption model found a spending *reduction*, demonstrates that the disagreement in this literature tracks scope and assumption choices far more than it tracks the ideological orientation of the institution producing the estimate. Where models converge in their scope and assumptions (comprehensive-benefit federal cost), estimates converge regardless of institutional orientation; where models diverge in scope and assumptions (deep payment cuts, narrower-benefit designs, optimistic administrative-savings and utilization assumptions), estimates diverge dramatically. This is direct evidence that the "same data, different conclusions" pattern common in this debate is substantially explained by assumption choices, not by data disputes or institutional bad faith.

Does the study model behavior realistically? This is the weakest point of several models. The Galvani/Yale model's assumption of zero utilization increase among the already-insured despite fully eliminating their cost-sharing directly contradicts one of the most robustly replicated findings in health economics (the original RAND Health Insurance Experiment of the 1970s–80s, and the Oregon Health Insurance Experiment more recently, both find utilization rises meaningfully when cost-sharing falls). CBO's models, by contrast, explicitly incorporate induced-demand elasticities and are more behaviorally realistic on this dimension, though even CBO acknowledges substantial uncertainty about exactly how large the induced-demand effect would be in a system as different from Oregon's Medicaid-only natural experiment as a comprehensive national program would be.

Does it account for increased utilization after universal coverage? As above: CBO yes (explicitly and centrally, with a documented $207–718 billion range); RAND's national model yes (with an explicit assumption that 50% of new demand goes unmet due to supply constraints — itself a debatable assumption, addressed below); Urban Institute yes (its 20.6% utilization-increase assumption, criticized by PNHP as too large); Galvani/Yale, only partially (accounts for increased utilization among the newly insured but not the already-insured, the specific limitation identified by independent fact-checkers).

Does it account for changes in provider payment? All major models address this directly, since it is the single largest driver of the bottom-line result (see above); the meaningful methodological differences are in the specific rate assumed, not whether the variable is modeled.

Does it account for pharmaceutical price reductions? Most do, with a wide range of assumed savings: Mercatus assumes $846 billion in ten-year drug savings from negotiation and generic substitution; Urban Institute assumes drug prices 25% below Medicare/50% above Medicaid; RAND's national estimate assumes drug prices 10% below current Medicare prices. These are all plausible in direction (given the well-documented, roughly 2.8-times US-versus-peer-country brand-drug price gap established by RAND's international comparison work) but vary substantially in assumed magnitude, and none of the models reviewed here directly incorporates the post-2022 Inflation Reduction Act Medicare drug-negotiation program's actual, now-observable early results as a real-world calibration point, since most of these models predate that program's implementation — a genuine gap in the currency of this literature that a future update should address.

Does it account for administrative savings? Yes, in every model, though — as emphasized above — this is the least contested and least consequential variable for the ultimate result, contrary to its outsized role in political messaging on both sides.

Does it account for changes in employer compensation? This is a substantial and under-addressed gap across nearly the entire literature reviewed. Standard labor economics holds that employer-sponsored health insurance is a form of compensation, and that relieving employers of premium obligations should, in a competitive labor market and over time, translate at least partially into higher wages — a dynamic the Heritage Foundation's static distributional analysis explicitly does not model, and that most of the cost-model literature reviewed here also does not model with much granularity. CBO's dynamic macro model (Working Paper 2022-02) is the most complete treatment found in this review, projecting private non-health consumption per capita rising roughly 11.5% by 2030 under its illustrative scenarios — implicitly incorporating some wage/compensation feedback — but even CBO flags substantial uncertainty about the speed and completeness of this pass-through in practice.

Does it account for taxes and household spending simultaneously? This is arguably the most poorly integrated question in the entire literature. Most cost models report either total national health spending or federal budget impact, but comparatively few directly integrate all four components simultaneously (federal taxes, state taxes, foregone employer premium contributions redirected to wages, and eliminated household premiums/out-of-pocket costs) into one "total household financial position" metric that would let a typical family compare their current all-in health costs to their all-in costs, including new taxes, under a specific single-payer design. CRFB's financing-package analysis comes closest by explicitly modeling the progressivity of various tax options relative to current household health spending, and its finding that most financing options are more progressive than the status quo is one of the more useful integrative findings in the literature, but even this analysis does not fully net out household premium and out-of-pocket savings against the specific new tax burden household by household.

Does it distinguish government spending from total national spending? The clearest and most consequential example of this distinction being lost, rather than made, in public discourse is the Mercatus/Blahous study, whose federal-budget estimate ($32.6 trillion) was repeatedly and, per Blahous's own public correction, incorrectly cited by some commentators as evidence about total national health spending falling. The studies that are most careful about this distinction (CBO, Urban Institute) report both figures separately and explicitly.

Does it use static or dynamic modeling? The Heritage Foundation analysis is explicitly static (no behavioral response modeled) and should be weighted accordingly, as its own limitations section implicitly acknowledges by not attempting dynamic modeling. CBO's Working Paper 2022-02 is the most fully dynamic model in this literature, using a general-equilibrium overlapping-generations framework. RAND's models are dynamic microsimulations but not full general-equilibrium macro models. This distinction matters enormously for any estimate involving a change of this scale (potentially 5–10% of GDP moving through different channels), since static models cannot capture wage pass-through, labor-supply responses, or the macroeconomic feedback of large tax or spending changes.

Does it account for supply constraints? This is one of the more genuinely uncertain and underexplored dimensions across the literature. RAND's national model explicitly assumes 50% of new demand goes unmet due to supply constraints — a strong assumption that, if wrong in either direction, would substantially change the utilization and cost estimates. CBO's methodology paper explicitly flags that demand growth would likely exceed supply growth, producing unmet demand in the range of roughly 1–6 percentage points depending on design, but treats this primarily as a coverage/access question (longer waits, some forgone care) rather than as a hard cost-ceiling; it is worth noting that no model reviewed here incorporates detailed, geographically specific data on current primary-care physician shortages, which are already documented in many parts of the U.S. under the status quo and would likely be exacerbated, at least during a transition period, by any coverage expansion.

Does it account for provider responses (e.g., reduced hours, early retirement, practice closure)? This is the least well-addressed question in the entire cost-modeling literature reviewed here. No model identified in this review incorporates a rigorous, data-driven estimate of how physicians, especially specialists facing the largest relative payment-rate cuts, would adjust hours, retirement timing, or practice location in response to a single-payer system's payment rates. This is a genuine and significant gap, not merely an omission of secondary

importance, since even modest reductions in physician labor supply in response to lower payment rates could materially worsen the supply-constraint problem identified above, particularly during a multi-year transition period.

Does it account for technological change? None of the major cost models reviewed here explicitly and rigorously models how a shift to single-payer financing, with different payment structures (e.g., global budgets or bundled payments as opposed to fee-for-service), might affect the rate of future medical-technology adoption and diffusion — a potentially significant long-run consideration given the well-documented finding (Section 7) that the U.S.'s fee-for-service-heavy system appears to reward faster adoption of some capital-intensive technologies. This should be treated as a highly uncertain, under-modeled dimension of the long-run single-payer question.

Does it adequately address transition costs? This is arguably the most consistently underexamined question across the entire literature reviewed. Every model reviewed here estimates a post-transition steady state (typically a ten-year window beginning at an assumed implementation date) rather than modeling the transition period itself — the multi-year process of unwinding existing insurance contracts, retraining or displacing an estimated 1.5–2 million private health-insurance-industry workers (a figure drawn from broader labor-statistics estimates of health-insurance-sector employment, not independently re-verified in this review's research process and flagged here as an estimate requiring separate verification), building new claims infrastructure, and managing the political and administrative complexity of a nationwide payment-system changeover comparable in scale to nothing the U.S. government has previously attempted. Vermont's 2014 experience — where anticipated administrative savings failed to materialize even at the planning stage, before implementation ever began — is the most direct empirical, if small-scale, evidence available on transition risk, and it points toward significant caution about assuming a smooth, on-schedule, on-budget transition at national scale.

Where apparently contradictory studies actually differ because of assumptions rather than empirical evidence

The clearest case in this literature is the Mercatus-versus-Urban-Institute "agreement despite opposite ideology" finding described above: once the benefit-design scope (comprehensive, LTSS-inclusive) is held constant, two institutions with opposite starting points reach similar federal-cost conclusions, which is strong evidence that the underlying empirical facts (provider payment levels, current spending baselines) are not seriously in dispute between these two institutions — only the framing and political use of the resulting numbers differs. Conversely, the sharp divergence between the Urban Institute and PNHP over administrative-savings estimates (9.5%→6% versus 9.5%→1.8%) is a genuine assumption dispute rooted in a real empirical disagreement about how much of the U.S.-Canada administrative-cost gap is attributable to the financing model itself versus to other structural differences (labor-market wage levels, state regulatory complexity, provider ownership structures) that would not automatically disappear under single payer — an empirical question that, per this review's assessment, remains genuinely unresolved and is treated as such in Section 6.

6. What We Know, What We Estimate, and What We Don't Know

High-confidence findings (supported by multiple independent sources, methodologically strong evidence)

- The United States spends far more on health care than any peer nation, in absolute terms, per capita, and as a share of GDP (CMS NHE data; OECD comparisons), and this gap has widened over the past two decades.
- This spending gap is driven predominantly by higher prices for the same underlying units of care — drugs, physician services, hospital procedures — not by higher utilization, which is often comparable to

or lower than peer-country averages (Anderson et al. 2003, 2019; Papanicolas et al. 2018 — independently replicated by two separate research teams over fifteen years).

- U.S. administrative overhead is substantially higher than in peer countries, on the order of two-and-a-half to four times, depending on measurement scope (Himmelstein/Woolhandler; Papanicolas et al.; Tseng et al. micro-costing) — the exact multiplier is contested, but the qualitative and directional finding is not, and is corroborated by researchers with no institutional stake in single-payer advocacy.
- U.S. brand-name pharmaceutical prices are roughly four times the average of comparator high-income countries, while U.S. generic prices are lower than most comparators (RAND 2021, 2024) — one of the most rigorously and repeatedly replicated findings in health economics.
- Nearly half of lower- and average-income Americans report skipping needed care due to cost, a rate far exceeding any peer country (Commonwealth Fund International Health Policy Survey) — a direct, minimally confounded measure of the U.S. access disadvantage.
- U.S. amenable/avoidable mortality (deaths preventable by timely, effective health care) is worse than in peer countries, and has been so consistently since this measure was first developed (Nolte & McKee 2008; Commonwealth Fund Mirror, Mirror series through 2024).
- U.S. maternal mortality is substantially higher than peer-country averages, and the within-U.S. racial disparity in this outcome is itself larger than the U.S.-versus-peer-country gap (Commonwealth Fund 2025).
- The U.S. performs comparably well or better than peer countries on several measures of clinical quality once care is actually delivered — cancer-specific survival (CONCORD-3), 30-day acute myocardial infarction and stroke mortality, and post-surgical patient safety indicators.
- Defensive medicine and malpractice costs, while real, are not a major driver of the U.S.-peer spending gap (Mello et al. 2010) — a well-quantified, minority finding that runs against a common political claim.
- Every credible cost model reviewed, across the ideological spectrum, agrees that single-payer financing would produce some administrative savings; the size of that saving, in isolation, is not the primary source of disagreement among credible analysts about whether total national health spending would rise or fall.
- Maryland's all-payer hospital rate-regulation experience demonstrates that administered pricing — a mechanism often assumed to require single-payer financing — is achievable within a multipayer system.

Moderate-confidence findings (credible evidence, but dependent on methodological or modeling assumptions)

- The specific magnitude of achievable administrative savings under a U.S. single-payer system is uncertain and contested, ranging across credible estimates from roughly 6% to as low as 1.8% of insurance-related spending, depending on assumptions about how much of the U.S.-Canada administrative-cost gap is attributable to financing structure versus other factors (labor markets, state regulatory complexity) that would not automatically disappear.
- Whether total U.S. national health spending would rise or fall under a specific single-payer design depends almost entirely on the provider-payment-rate assumption built into that design; the credible range spans from a meaningful decrease (models assuming Medicare-level or below rates) to a meaningful increase (models assuming rates closer to current private-insurance levels), and this is a genuinely uncertain empirical and political question, not a settled one.
- Federal government spending would increase substantially under any comprehensive, LTSS-inclusive single-payer design — credible estimates across ideologically opposed institutions (Mercatus, Urban Institute) converge on a ten-year federal cost in the range of $32–33 trillion for such a design, a

genuinely well-corroborated moderate-confidence figure, though it is specific to that comprehensive benefit design and would be substantially lower for a narrower benefit package.

- Health care utilization would likely increase following the elimination of cost-sharing and the extension of coverage to the currently uninsured, consistent with well-established health-economics findings on induced demand (RAND Health Insurance Experiment, Oregon HIE); the magnitude of this increase, and the degree to which existing provider supply could absorb it without meaningful new waiting times or access friction, is uncertain.
- Health outcomes on the access- and equity-sensitive measures identified above (unmet need, medical debt, maternal mortality disparities, amenable mortality) would plausibly improve under expanded, more affordable coverage, based on the causal and quasi-experimental evidence reviewed (Oregon HIE, Taiwan's 1995 natural experiment, Massachusetts reform); the magnitude of improvement at a national scale, and the timeline over which it would materialize, cannot be estimated with high precision from the available natural experiments, which were smaller in scale and scope than a national single-payer transition would be.
- Provider payment reductions of the magnitude assumed in several favorable single-payer cost models (yielding net national savings) may not be economically or politically sustainable at the scale and speed assumed, based on the qualitative pattern observed in Vermont's abandoned 2014 implementation attempt and on the fact that even a critic of single-payer costing (Blahous) and a comprehensive-benefit-design study (Urban Institute) both flag the underlying payment-rate assumption as a key point of uncertainty.

Highly uncertain findings (areas where reasonable analysts reach substantially different conclusions)

- Whether the net effect on total national health spending, integrating administrative savings, provider payment changes, induced demand, and drug-price reductions simultaneously, would be positive or negative for any specific real (as opposed to illustrative) legislative proposal. This is genuinely and legitimately contested among credible, non-fringe analysts, and the disagreement traces to differing, defensible-in-isolation assumptions about provider behavior and payment sustainability, not to a resolved empirical fact that one side is simply ignoring.
- The transition costs and risks of a national-scale single-payer implementation — including the speed and completeness of employer-to-wage compensation pass-through, the labor-market effects of displacing an estimated one to two million private health-insurance-industry employees, and the operational risk of building new national claims infrastructure — are almost entirely unmodeled in the quantitative literature reviewed and should be considered a substantial, unquantified source of uncertainty rather than a solved problem.
- Whether existing U.S. health-care provider supply (physicians, especially primary-care physicians and certain specialties, along with hospital capacity in underserved regions) could absorb the induced demand from universal, cost-sharing-free coverage without meaningful new access friction (waits, geographic rationing) is genuinely unresolved, and no model reviewed here incorporates the kind of detailed, geographically granular supply data that would be needed to answer it with confidence.
- The degree to which U.S. technological innovation and adoption speed (in pharmaceuticals specifically, and plausibly in medical devices, though this is less well-evidenced) would be affected by a shift away from the current fee-for-service, multipayer payment structure is genuinely uncertain, understudied in the empirical literature reviewed here, and a legitimate point of concern that credible analysts on both sides treat very differently depending on their priors about the relationship between payment structure and innovation incentives.

- How much of the U.S. life-expectancy and mortality disadvantage relative to peer countries is attributable to the health-care financing system specifically, as opposed to behavioral, social, and demographic factors outside that system, is explicitly and appropriately left without a precise quantitative answer by the National Academies' own 2013 panel — a genuinely open scientific question, not merely a political one.

Claims frequently made in political debate that are not adequately supported by the evidence reviewed here

- The claim, sometimes made by single-payer advocates, that Medicare for All would straightforwardly "save trillions" in aggregate national health spending with high confidence is not well supported; the credible range spans from meaningful savings to meaningful cost increases depending on provider-payment assumptions that are themselves uncertain and contested, and the specific studies most often cited for large savings (Galvani et al.) have documented, independently identified methodological weaknesses (excluded LTSS costs, an implausible zero-utilization-response assumption).
- The claim, sometimes made by single-payer opponents, that the administrative-savings component of single-payer cost models is fabricated or trivial is not well supported; independent, non-advocacy-affiliated research (Papanicolas et al.; Tseng et al.) corroborates the qualitative finding of substantial U.S. administrative overhead, even though the specific magnitude of achievable savings is genuinely contested.
- The claim that other countries' lower spending and better outcomes prove that single-payer financing specifically, as opposed to price regulation, administrative simplification, or universal-coverage mandates more broadly, is the necessary and sufficient cause of those countries' comparative performance is not well supported; the mechanism-attribution analysis in Section 7 shows that several of the most important cost-saving mechanisms (administered pricing, drug-price regulation) are separable policy choices achievable within a multipayer system, as Germany, the Netherlands, and Maryland's own domestic experience demonstrate.
- The claim that the U.S. outcomes disadvantage relative to peer countries reflects a uniformly lower-quality health-care delivery system is not well supported; the U.S. performs comparably or better than peers on several clinical-quality measures (cancer survival, acute cardiovascular mortality, patient safety) even while badly underperforming on access- and equity-sensitive measures — a pattern inconsistent with a simple "worse care" narrative and more consistent with an "unequal access to good care" narrative.
- The claim that malpractice reform or "defensive medicine" elimination represents a major lever for U.S. cost control is not well supported by the most rigorous quantitative study of this specific question, which found these costs "unlikely to be a source of significant savings" at the system level.
- The claim that Medicare for All would necessarily require large middle-class tax increases with no offsetting savings, framed as a straightforward net loss for typical households, is not well supported without qualification; the most rigorous financing-options analysis available (CRFB) found most plausible financing packages more progressive than the current system, and static distributional analyses that reach starker conclusions (Heritage Foundation) do not model the wage pass-through from eliminated employer premium obligations that standard labor economics predicts would at least partially offset new tax burdens for most households.

7. International Evidence: Examining Why, Not Just That

It is not enough to observe that other high-income countries spend less and, on many measures, achieve better outcomes; the review must examine the mechanisms and ask which are attributable to the financing model itself and which are separable policy choices that a multipayer system could, in principle, also adopt.

Healthcare prices. The foundational and repeatedly replicated finding (Anderson et al. 2003, 2019; Papanicolas et al. 2018) is that U.S. prices for the same underlying units of care — physician visits, procedures, hospital days, drugs — are dramatically higher than in peer countries, while utilization is often comparable or lower. This is the single most robust finding in the entire international-comparison literature, replicated across three independent research efforts spanning fifteen years. **Attribution:** largely a separable policy choice about bargaining structure and regulation, not an intrinsic feature of single- versus multi-payer financing. Germany, Japan, and Switzerland — all multipayer systems — achieve U.S.-peer-comparable low prices through all-payer rate regulation or binding fee schedules negotiated across all insurers simultaneously, without collapsing to a single payer.

Physician compensation. U.S. general-practitioner salaries are roughly double the peer-country average (Papanicolas et al. 2018; Laugesen & Glied 2011), and the gap is even larger for procedural specialties. **Attribution:** a separable policy choice tied to fee-schedule design (the U.S. RBRVS/RUC valuation process, criticized by health economists across the spectrum for systematically undervaluing primary care relative to procedures) and residency-slot allocation via graduate medical education funding, not an inherent feature of the payer structure.

Hospital payment. The clearest domestic proof that administered pricing is separable from financing-model choice is Maryland's all-payer hospital rate-setting system, the only such system in the United States, which retains private insurance, Medicare, Medicaid, and employer coverage while imposing state-regulated global budgets on hospitals. Rigorous difference-in-differences evaluation (Morrison et al. 2021) found meaningfully slower Medicare and commercial spending growth under this model relative to comparison states. **Attribution:** directly separable — Maryland demonstrates a multipayer system can adopt single-payer-style administered pricing for hospitals specifically, without changing who pays.

Pharmaceutical prices. U.S. brand-name drug prices are roughly 4.2 times the average of 33 comparator countries (RAND 2024); generics, by contrast, are cheaper in the U.S. than in most comparators. **Attribution:** separable — every peer country with a private-insurance-based system (Germany, Switzerland, Netherlands) achieves lower drug prices through government reference pricing or health-technology-assessment-linked pricing thresholds, mechanisms a multipayer system can adopt (as the U.S. partially did, for Medicare specifically, through the 2022 Inflation Reduction Act) without becoming single-payer.

Insurance administration. This is the mechanism most directly and mechanically tied to the financing model itself. Multi-payer billing complexity — hundreds of different insurers, each with different claims formats, prior-authorization rules, and eligibility-verification processes — imposes real, well-documented administrative costs on both payers and providers (Section 2.1). **Attribution:** partially inherent to multipayer fragmentation, but not entirely — Germany and the Netherlands, both multipayer systems, keep administrative costs far below U.S. levels through standardized, uniform claims and billing rules imposed across all insurers, showing that simplification and standardization can substantially mitigate (though perhaps not fully eliminate) this cost without full consolidation to a single payer.

Negotiating power / provider and payer market concentration. U.S. hospital markets have consolidated substantially, and market power translates directly into higher prices (Cooper, Craig, Gaynor & Van Reenen 2019): monopoly hospitals charge 12–15% more than hospitals in more competitive markets. **Attribution:** a separable competition-policy problem (antitrust enforcement, certificate-of-need law reform) reflecting weak *countervailing* payer power against consolidated providers within the current fragmented multipayer system; administered pricing (whether via a single payer or an all-payer regulatory body, as in Maryland) directly neutralizes this specific pathology without requiring a change in who pays.

Government regulation / price-setting mechanisms abroad. Global budgets, all-payer rate setting, and reference pricing are all mechanisms available to and used by non-single-payer systems abroad, as established by the Maryland and pharmaceutical-pricing evidence above. **Attribution:** a policy choice about bargaining structure and regulatory design, separable from the single-payer/multipayer axis.

Healthcare utilization. Contrary to a common assumption in political debate, U.S. utilization is not uniformly higher than peer countries; it runs in both directions depending on the specific service (Peterson-KFF Health System Tracker 2024–2025). The U.S. has roughly half as many in-person physician visits per capita as the peer average and shorter average hospital stays, but more hospital discharges per capita and comparable-to-higher diagnostic imaging volumes. Coronary bypass surgery, for instance, costs 2.9 times the peer-nation Medicare-equivalent average in the U.S. despite comparable underlying volume — a pure-price effect. **Attribution:** utilization patterns more plausibly reflect care-delivery norms (discharge practices, fee-for-service incentives that reward discrete billable encounters and imaging) than the financing model per se, though fee-for-service payment design — common in, but not exclusive to, the U.S. multipayer system — plausibly does incentivize certain utilization patterns (more discrete billable procedures, more imaging) that a global-budget or capitated payment system, whether single-payer or not, would tend to reduce.

Primary care versus specialist access. A real, though frequently overstated, U.S. generalist shortage exists relative to peer countries after correcting for a common definitional mismatch in how "generalist" is counted (Laugesen 2018). **Attribution:** separable — driven by GME residency-slot funding allocation and relative fee-schedule valuation (favoring procedures over cognitive/primary care), which a multipayer system could reform without changing its financing structure.

Technology adoption. The U.S. is the fastest-launching major pharmaceutical market in the world, with over half of new drugs launching first in the U.S. (RAND Health Quarterly 2024) — a genuine U.S. system strength largely attributable to FDA review speed and the absence of a government price/health-technology-assessment gate delaying launch, rather than to the multipayer financing structure per se. No comparably rigorous cross-national evidence on device, imaging, or surgical-technology diffusion speed was identified in this review's research process, and this should be treated as an open question rather than assumed to follow the same pattern as pharmaceuticals. To the extent U.S. fee-for-service payment does reward capital investment in billable imaging and procedural technology more than global-budget systems abroad, this dimension is plausibly, though not conclusively, tied to payment-model design generally rather than the single-payer/multipayer distinction specifically.

Defensive medicine. Malpractice and defensive-medicine costs, while real (an estimated $45.6 billion per year in defensive medicine specifically — Mello et al. 2010), are explicitly found by the authors of the most rigorous study on this question to be "unlikely to be a source of significant savings" at the system level — accounting for only about 2.4% of total U.S. health spending for the liability system as a whole. **Attribution:** a U.S. tort-law feature, separable from the insurance-financing model; different malpractice regimes abroad (no-fault compensation, damage caps) demonstrate this is a legal-system policy lever, unrelated to the payer structure.

Social determinants of health, population demographics, and public health infrastructure. The National Academies' 2013 panel explicitly and carefully declines to offer a single clean percentage-based attribution of the U.S. mortality/life-expectancy disadvantage across behavioral, social, environmental, and health-system factors, emphasizing their interconnection; Preston and Ho's complementary analysis argues that U.S. cancer-screening and survival performance suggests the health-care *delivery system* itself is not obviously underperforming relative to peers, pointing instead toward behavioral and social factors (obesity, historical smoking rates, homicide, drug overdose) as larger drivers of the U.S. life-expectancy gap specifically, as distinct from the spending gap. **Attribution:** these factors are, definitionally, not attributable to the health-insurance financing model, and this review treats the National Academies' explicit caution against forcing a false precision here as the appropriate standard.

Bottom line on mechanism attribution

Sorting the evidence above by attribution yields a genuinely useful distinction. **Most directly and structurally tied to the financing model:** insurance administrative complexity (though even here, mitigation short of full single-payer consolidation is demonstrably possible, as Germany and the Netherlands show). **Separable policy**

choices, achievable with or without single payer: provider price regulation (Maryland proves this), pharmaceutical reference pricing, primary-care/specialist fee-schedule rebalancing, antitrust/market-concentration remedies, and malpractice reform. **Not plausibly attributable to the financing model at all:** the large majority of the U.S. life-expectancy and mortality disadvantage relative to peer countries, which the National Academies and related literature attribute substantially to behavioral, social, and demographic factors outside the health-care system's control. This sorting matters directly for Section 8's evaluation of single payer specifically: it implies that adopting single-payer financing would plausibly capture the administrative-savings mechanism most directly, could capture some of the pricing-mechanism benefits (though these are also achievable through narrower reforms within the current multipayer structure), and would likely have only a limited direct effect on the largest share of the U.S. outcomes disadvantage, which stems from factors outside the health-care financing system altogether — while very plausibly improving the specific, well-documented subset of outcomes most directly tied to access and affordability (unmet need, medical debt, amenable mortality, maternal mortality disparities).

8. Single Payer Specifically: What the Evidence Implies

This section addresses each of the requested questions directly, distinguishing throughout between single-payer financing as an abstract mechanism and the specific bundled Medicare for All legislative proposals that most U.S. quantitative cost estimates actually model (see Section 1).

Could single payer reduce total national healthcare spending? Possibly, but this is not a settled or high-confidence conclusion — it depends almost entirely on provider payment rates and the degree to which induced demand and administrative savings offset each other, per the assumption analysis in Section 5. Designs assuming payment rates near or below current Medicare levels tend to show net national savings (CBO's lower-payment illustrative scenario, RAND's national model showing modest net savings only under stringent unmet-demand assumptions); designs assuming payment rates closer to current average multipayer rates tend to show net cost increases even after accounting for administrative savings (CBO's higher-payment scenario; several of the Urban Institute/Mercatus benefit-inclusive models). There is no credible basis in this literature for stating with high confidence, in either direction, what the net national-spending effect of a specific real proposal would be without specifying its payment-rate design.

Would it increase federal government spending? Yes, with high confidence, for any comprehensive, LTSS-inclusive design of the kind most Medicare for All legislation has proposed — this is one of the more robust findings in the literature, corroborated by ideologically opposed institutions (Mercatus and Urban Institute both around $32–33 trillion over ten years). This would represent a large fiscal shift regardless of whether total national spending rises or falls, because it converts currently private spending (employer premiums, household premiums and out-of-pocket costs) into public spending, which must be financed through some combination of new taxes, reduced other federal spending, or debt.

Would it reduce household and employer spending? Households would very likely see reduced direct premium, deductible, and out-of-pocket costs, given the near-elimination of cost-sharing in most Medicare for All designs and the current substantial magnitude of those costs ($9,325–$26,993 in average annual premiums, plus deductibles and out-of-pocket spending, per the 2025 KFF Employer Health Benefits Survey). Whether households come out ahead on a fully netted basis (accounting for new taxes) depends on the specific financing mechanism chosen and on income level; CRFB's analysis found most financing options more progressive than current arrangements, implying many lower- and middle-income households would likely see net financial benefit while higher-income households would more likely see net cost increases, though this varies by the specific financing mix and is not uniform across every proposed design. Employers currently providing health benefits would be relieved of a substantial direct cost (average $9,325–$26,993 per covered employee annually before worker contributions), and standard labor economics predicts at least partial pass-through of this savings into wages over time, though the completeness and speed of that pass-through is uncertain and contested, and is essentially unmodeled in most of the specific cost studies reviewed here.

How much administrative savings are realistically achievable? This is genuinely uncertain, with credible estimates ranging from CBO's roughly 1.5–2% of total spending to the PNHP-affiliated estimate of savings toward Canadian levels (1.8% of a narrower insurance-overhead base, implying larger dollar totals when combined with provider-side savings). The most defensible position, given the evidence reviewed, is that meaningful administrative savings are highly likely to occur (this is essentially undisputed across the ideological spectrum), but that the largest, most optimistic estimates (approaching full Canadian-level convergence) likely overstate how quickly and completely U.S.-specific structural factors (labor-market wage levels, existing health-IT infrastructure, state regulatory complexity, and transition costs during implementation) would allow that convergence to occur.

Could pharmaceutical prices be substantially reduced? Yes, with reasonably high confidence in direction, given the well-documented roughly four-times gap between U.S. brand-name drug prices and the peer-country average (RAND 2024) and the demonstrated effectiveness of government price negotiation and reference-pricing mechanisms abroad. The magnitude of achievable savings under a specific single-payer design is less certain and depends on the negotiating mechanism's design and on pharmaceutical manufacturers' pricing and R&D responses, neither of which is fully modeled in the literature reviewed here; it is also worth noting that some drug-price reduction is achievable, and has already begun to be achieved, within the current multipayer system (the Inflation Reduction Act's Medicare negotiation provisions), meaning this benefit is not exclusively available through single-payer financing.

Would provider payment reductions be necessary? For a single-payer system to produce net national savings (as opposed to merely shifting spending from private to public sources), some degree of provider payment reduction relative to current average multipayer rates is very likely necessary, given that provider payment rates are the dominant driver of the bottom-line spending result across every model reviewed. The size of reduction that would be both fiscally meaningful and economically/politically sustainable (without triggering the kind of provider-supply response discussed below) is genuinely uncertain and is the central unresolved empirical and political question in this entire literature.

Would healthcare utilization increase? Yes, with high confidence, consistent with the well-established health-economics literature on induced demand following expanded coverage and reduced cost-sharing (RAND Health Insurance Experiment; Oregon HIE). The magnitude at a national scale, given the size and heterogeneity of the currently underinsured and uninsured U.S. population, is harder to estimate precisely than in either of those smaller natural experiments.

Would there be capacity constraints? This is a genuine, underexplored risk rather than a settled question. CBO's methodology explicitly anticipates unmet demand in the range of roughly 1–6 percentage points depending on design; RAND's national model assumes 50% of new demand goes unmet due to supply constraints. Neither of these figures should be read as a precise forecast, but both indicate that credible, non-advocacy-affiliated analysts across the spectrum expect some degree of access friction (longer waits, geographic rationing, or both) during a transition period, particularly given already-documented primary-care shortages in many parts of the United States under the status quo.

Would access improve or deteriorate? Both, in different senses, and probably simultaneously for different populations — this is one of the more nuanced findings this review can offer. For the currently uninsured and underinsured (a substantial population, given that nearly half of lower- and average-income Americans report cost-related access barriers), access would very plausibly improve, consistent with the causal and quasi-experimental evidence on coverage expansion (Oregon HIE, Taiwan 1995, Massachusetts). For those currently well-insured with ready access to a broad network of specialists and elective procedures, some deterioration in speed of access (longer waits for certain elective services) is a plausible, if uncertain, risk during any period in which demand growth outpaces supply growth, a pattern observed to varying degrees in most (though not all) single-payer and NHS-model countries abroad for certain elective procedures specifically, while remaining a genuinely open question, not settled by U.S.-specific data, in the U.S. context.

Would health outcomes likely improve? On the access- and equity-sensitive measures most directly and causally linked to coverage and affordability (unmet need, medical debt, financial hardship, and plausibly maternal and amenable mortality, per the Taiwan and Massachusetts evidence), yes, with moderate-to-high confidence. On the broader life-expectancy and overall-mortality measures most heavily influenced by behavioral, social, and demographic factors outside the health-care system (per the National Academies' 2013 analysis), the effect would likely be smaller and slower to materialize than advocates sometimes suggest, since a large share of the U.S. mortality disadvantage is not primarily a function of insurance coverage or financing structure.

Distributional effects across income groups. Lower- and middle-income households, who currently bear a disproportionate share of cost-related access barriers and medical debt and who would benefit most from progressive financing replacing regressive premium and out-of-pocket structures, would very plausibly see the largest net benefit; higher-income households would more likely see a net cost increase under most progressive financing designs, consistent with CRFB's financing-options analysis.

Effects on employers. Employers currently offering coverage would be relieved of a substantial direct cost, likely offset in part by new payroll-based taxes depending on financing design; net effects would vary considerably by industry, firm size, and current benefit generosity, and are not uniformly modeled across the literature reviewed.

Effects on hospitals. Hospitals would very likely see payment-rate changes (a mix of upward pressure from previously uninsured/uncompensated-care patients now being reimbursed, and downward pressure from privately-insured patients being shifted to lower, Medicare-like rates); net effects would vary substantially by hospital payer mix, with hospitals currently serving a larger share of privately insured patients facing more downward pressure and safety-net hospitals currently absorbing large uncompensated-care burdens potentially benefiting. This heterogeneity is acknowledged in the literature but not modeled hospital-by-hospital in any study reviewed here.

Effects on physicians. Primary-care physicians, whose current fees are closer to any plausible single-payer payment rate, would likely see smaller relative reductions (or in some designs, increases, if primary-care fee schedules were specifically rebalanced upward); proceduralist specialists, whose current private-insurance fees substantially exceed Medicare rates, would likely see the largest relative reductions — a dynamic that several credible analysts (across the ideological spectrum) flag as a potential driver of physician-supply responses (reduced hours, earlier retirement, specialty-choice shifts) that is not well quantified in the literature reviewed here.

Effects on pharmaceutical manufacturers. Manufacturers would very likely face reduced U.S. revenue under any design incorporating meaningful price negotiation or reference pricing, given that the U.S. currently accounts for a disproportionate share of global pharmaceutical profit relative to its share of prescription volume (RAND 2024 finding that the U.S. is 62% of OECD drug sales value but only 24% of volume). The knock-on effect on pharmaceutical R&D investment and the pace of new-drug development is a long-standing, genuinely contested question in health economics that this review's research process did not find a rigorous, current, non-advocacy-affiliated quantitative answer to, and which should be treated as a significant open uncertainty rather than a resolved one on either side.

Effects on insurers. The private health-insurance industry, including its roughly one-to-two-million-person workforce (an order-of-magnitude estimate, not independently verified in this review's research process), would face existential disruption under any Medicare for All design that eliminates the industry's role for covered benefits; this is among the most significant and least-modeled transition costs in the entire literature, encompassing both the human cost of large-scale job displacement and the practical challenge of building replacement claims and eligibility infrastructure at national scale within a defined implementation timeline.

Effects on state governments. States would very likely see both fiscal relief (elimination of the state share of Medicaid financing, a substantial current state budget item) and new administrative and, depending on design, financing obligations; a national single-payer program would also resolve the specific interstate risk-pool and adverse-selection problems that undermined the Vermont and (in a different way) the proposed New York state-level plans, since a national program does not face the same interstate competitive and migration pressures a single state does.

Effects on federal finances. As above, federal spending would increase substantially in dollar terms for any comprehensive design, requiring new financing (taxes, reduced other spending, or debt) of a scale roughly comparable to, or in the higher estimates exceeding, the entire current federal discretionary budget — a fiscal undertaking whose scale should not be understated regardless of one's view on its desirability.

Transition risks. This is the least well-quantified, most consistently underexamined dimension across the entire literature reviewed, encompassing labor-market disruption in the insurance industry, the operational risk of building new national claims and eligibility infrastructure within a defined timeline, the risk (illustrated at small scale by Vermont's 2014 experience) that anticipated savings fail to materialize even at the planning stage, and the political and administrative complexity of a payment-system changeover with no close precedent in U.S. government history. This review treats transition risk as a major, genuinely open source of uncertainty that the existing quantitative cost-modeling literature — which almost universally models a post-transition steady state rather than the transition itself — does not adequately address, and any credible discussion of single payer's likely real-world implications should treat this gap as a significant qualification on all of the steady-state cost and outcome estimates discussed above.

9. Comparing Competing Policy Scenarios

Scenario	Expected cost	Expected outcomes	Coverage	Administrative complexity	Political feasibility (2026)	Transition difficulty	Major risks	Major uncertainties
1. Current U.S. system	Highest per-capita/GDP share among high-income countries ($5.3T/18.0% GDP in 2024; projected ~$9.0T/20.6% GDP by 2034, CMS)	Best-in-class on some clinical-quality measures (cancer survival, acute cardiovascular mortality); worst-in-class among peers on access, equity, amenable mortality, maternal mortality, financial hardship	~26.6M uninsured in 2025, CBO projects rising toward ~37M by the mid-2030s absent policy change, due to expiring ACA subsidies and 2025 reconciliation-law effects	High — administrative costs ~8-34% of spending depending on measurement scope (Papanicolas et al.; Himmelstein et al.), well above any peer country	N/A (status quo, but not static — currently trending toward coverage losses)	N/A	Continued erosion of coverage as subsidies expire; continued price and cost growth outpacing wages; growing medical debt and disparities	How much further coverage erosion actually occurs depends on near-term legislative choices not yet finalized as of this review
2. ACA-based reforms (extend/expand subsidies, strengthen marketplace)	Modest incremental federal cost to extend enhanced subsidies (CRFB, Commonwealth Fund); far smaller fiscal footprint than single-payer designs	Would prevent projected coverage losses; unlikely to meaningfully close the access/equity gap identified in Sections 2 and 6 given continued underlying cost-sharing and network limitations	Would preserve roughly current coverage levels (~92-93% insured) rather than the erosion CBO now projects; would not achieve full universal coverage	Unchanged — retains full multipayer administrative complexity	Highest of any expansion option in the current Congress, given it extends existing, previously bipartisan-adjacent structures rather than creating new ones	Low — no structural change required	Continued exposure to periodic subsidy-expiration political fights (as in 2025-26); does not address underlying price or administrative-cost drivers	Whether Congress extends subsidies permanently or continues short-term extensions with recurring uncertainty
3. Public option	Premium and federal-subsidy savings depend almost entirely on how far below private-market rates the public plan's payment rates are set (RAND 2020; Urban Institute 2021) — a "plain vanilla" design with no rate leverage produces only modest savings	Modest coverage gains (RAND: 1.2-2.8 million fewer uninsured depending on design)	Partial — does not achieve universal coverage; effects concentrated in the individual/nongroup market	Moderate — adds one more payer option rather than simplifying the existing multipayer landscape	Moderate — has periodically had majority Democratic support but has not passed Congress; more incremental than single payer	Moderate — requires building a new public plan option and setting its payment rates, but does not require dismantling existing coverage arrangements	If payment rates are set too close to private rates, produces limited benefit; if set much lower, risks adverse effects on provider participation/access, an effect not fully modeled in the studies reviewed	How aggressively any enacted public option would set payment rates, and how providers would respond
4. Regulated multipayer universal coverage (Netherlands/Switz	No U.S.-specific quantitative model of this exact scenario was	Internationally associated with strong outcomes and near-universal	Would achieve near-universal coverage via mandate and	Would still require simplification/standardization of billing and claims across	Low in the current U.S. political environment — this model is not the	High — would require an individual mandate (politically	Mandate enforcement and political durability, given the fate of the	Whether a politically viable, enforceable individual-mandate

Scenario	Expected cost	Expected outcomes	Coverage	Administrative complexity	Political feasibility (2026)	Transition difficulty	Major risks	Major uncertainties
erland-style mandate + regulation)	identified in this review's research process; internationally, such systems achieve substantially lower spending shares of GDP than the U.S. through mandate, regulation, and administered pricing rather than single-payer consolidation	access, though cross-national comparison is subject to the same confounding caveats noted in Section 7	subsidy, similar in structural logic to the ACA's individual mandate (since repealed at the federal penalty level)	insurers to approach the administrative-cost performance of the Netherlands/Switzerland, per Section 7's analysis	subject of active federal legislative proposals as of this review, unlike the public option, Medicare buy-in, or Medicare for All	contested since the ACA-era mandate-penalty repeal) and much stronger insurance-market regulation than the U.S. currently has	ACA's own mandate; achieving Netherlands/Switzerland-level administrative efficiency would require substantial new standardization mandates on private insurers	mechanism could be re-established in the current U.S. political environment
5. Medicare buy-in / age-lowering (e.g., to 60)	$155 billion added federal deficit over 2026-2031 for age-60 buy-in (CBO 2022)	Primarily a coverage-source shift (from employer/Medicaid /nongroup to Medicare) for near-elderly adults who mostly already have some coverage, rather than a large net reduction in the uninsured	Minimal net reduction in uninsured (CBO: only ~0.4 million fewer uninsured out of 15.6 million newly eligible)	Adds Medicare enrollment complexity but does not restructure the broader multipayer system	Moderate — has recurring Democratic support (raised in 2020-21 debates) but has not passed; more incremental and narrower in scope than a public option	Low-moderate — extends an existing, well-established program rather than creating a new mechanism	Primarily a fiscal cost with limited coverage benefit relative to its cost, per CBO's own findings	Take-up rate assumptions (CBO assumed ~87% enrollment among the newly eligible)
6. Single payer / Medicare for All	Highly uncertain net national-spending effect (CBO range: -$743B to +$290B by 2030 across illustrative designs); federal spending increase of ~$32-33 trillion over ten years for a comprehensive, LTSS-inclusive design (Mercatus, Urban Institute convergence)	Plausible improvement on access- and equity-sensitive measures (unmet need, medical debt, maternal/amenable mortality); uncertain and likely smaller effect on overall life expectancy given the substantial behavioral/social component of the U.S. mortality disadvantage	Universal, by design	Would substantially reduce insurance administrative complexity (converging toward, though probably not fully reaching, Canadian levels per Section 5's assumption analysis), though this savings is the least consequential of the major disputed assumptions for the net-cost question	Lowest among all options in the current Congress and in recent state-level attempts (California's AB1900 failed a third time in 2026; Vermont abandoned its plan in 2014; Colorado's ballot measure was defeated 79%-21% in 2016)	Highest of any option — requires unwinding existing employer-based coverage for roughly half the population, building new national claims infrastructure, and managing large-scale insurance-industry labor displacement, none of which is well-modeled in the existing quantitative literature (Section	Provider-payment-rate sustainability, provider-supply response, transition execution risk, pharmaceutical-innovation effects, and the risk (illustrated by Vermont) that anticipated savings fail to materialize even at the planning stage	Nearly every major cost-and-outcome parameter identified in Sections 5, 6, and 8 remains genuinely and legitimately contested among credible analysts

Scenario	Expected cost	Expected outcomes	Coverage	Administrative complexity	Political feasibility (2026)	Transition difficulty	Major risks	Major uncertainties
						5)		

Observations from the comparison

The six scenarios above form a rough continuum of increasing coverage ambition and increasing transition difficulty, but this review's evidence base does not support a claim that this continuum also represents a continuum of increasing net financial benefit — the relationship between coverage ambition and net cost is contingent on the specific payment-rate and financing design chosen within each scenario, not a simple linear relationship. Scenario 2 (ACA-based reforms) is the most fiscally modest and most politically feasible in the immediate term, but does the least to address the underlying price and administrative-cost drivers identified throughout this review. Scenario 6 (single payer) has the largest plausible upside on the access- and equity-sensitive outcome measures this review finds best-supported by evidence, and the largest plausible administrative-savings capture, but also carries by far the largest transition risk and the least political feasibility of any option reviewed, and its net financial effect on total national spending is genuinely uncertain rather than reliably favorable. Scenarios 3–5 (public option, regulated multipayer, Medicare buy-in) occupy an underexplored middle ground in the U.S. quantitative literature — comparatively modest in ambition, cost, and transition risk, but also modest in expected coverage and outcome gains relative to the scale of the access and equity problems this review documents in Sections 2 and 6.

10. Why Do Credible Analysts Disagree?

This section applies the specific diagnostic categories requested to the major documented disagreements identified throughout this review, to determine whether apparent disagreement reflects a real empirical dispute, a scope/definition difference, or a methodology choice.

Mercatus ($32.6T federal, 2018) vs. Urban Institute ($32.0T federal, 2016). Not a real disagreement once methodology is examined — these two ideologically opposed institutions converge closely on the federal-cost estimate for a comparable, comprehensive, LTSS-inclusive benefit design. The remaining difference is attributable to modest differences in assumed administrative-savings rates and drug-price assumptions, not to a fundamental empirical dispute. This convergence is, on its own, one of the more evidentially important findings in this entire review, because it demonstrates that when scope and core assumptions are held roughly constant, institutional ideology does not drive the result.

Mercatus/Urban Institute ($32-33T federal cost, comprehensive design) vs. Galvani et al. ($450B/year net national savings) vs. PERI (-9.6% national spending). A real disagreement, but one driven overwhelmingly by **scope and utilization/payment-rate assumptions**, not by a dispute over the underlying spending data. Galvani's model excludes roughly $4 trillion in long-term-care costs the actual Sanders legislation would cover, assumes hospitals accept Medicare-level payment as a system-wide ceiling, and assumes zero utilization increase among the already-insured — three assumption choices, each independently identified as questionable by non-partisan fact-checkers, that together produce a far more favorable result than the comprehensive-benefit-design models. PERI's more favorable estimate similarly reflects more aggressive administrative and drug-price-negotiation savings assumptions than the comprehensive-design models use. This is a case where the disagreement is **methodological/assumption-driven, not data-driven** — the underlying CMS spending baseline all these models start from is the same.

CBO's range of -$743B to +$290B in national spending by 2030 across its own five illustrative scenarios. This "disagreement," notably, comes from a single institution modeling the sensitivity of its own results — a direct demonstration that **provider-payment-rate and cost-sharing assumptions**, not institutional bias or differing data, are sufficient on their own to span a range wider than the difference between many of the competing external studies. This is arguably the single most important piece of evidence in the entire review for the proposition that most of the public "disagreement" over single-payer costing is a disagreement over policy-design choices dressed up as a disagreement over facts.

Urban Institute (assumes ~zero net provider-side administrative savings) vs. PNHP/Himmelstein-Woolhandler (assumes provider-side savings up to Canadian levels, ~$2.57 trillion over ten years). This is the clearest

example in the literature of a **genuine, unresolved empirical disagreement** rather than a scope or definitional difference. Both sides agree the U.S. has higher provider-side administrative costs than Canada (the underlying data are not disputed). They disagree about how much of that gap is caused by the multipayer financing structure itself (and would therefore disappear under single payer) versus how much is caused by other structural features of the U.S. health system — labor-market wage levels, more complex hospital ownership and regulatory structures, state-by-state licensure, existing health-IT architecture — that would not automatically change even if the payer structure did. This is a real, evidence-relevant empirical question that the literature reviewed here does not resolve, and it should be presented to readers as **genuinely open**, not as a case where one side is simply correct.

Heritage Foundation (73.5% of Americans worse off) vs. CRFB (most financing options more progressive than current system). Primarily a **methodology** difference: Heritage's analysis is static (no behavioral response modeled, single financing mechanism assumed) while CRFB's analysis incorporates dynamic macroeconomic modeling (Penn Wharton) across multiple financing options and explicitly compares progressivity to the current system. This is not a data dispute — it is a dispute over whether a static or dynamic modeling approach, and a single financing assumption versus a range of options, is the more appropriate and informative way to answer the underlying question. This review treats the dynamic, multi-option approach as more informative, consistent with the general finding (Section 5) that static models cannot capture the wage and behavioral effects that are large relative to the direct effects being measured in a policy change of this scale.

Cross-cutting pattern. Sorting these disagreements by the categories requested: **utilization assumptions** and **provider payment assumptions** together explain the large majority of variance across the "headline number" cost estimates (CBO's own scenario range demonstrates this most cleanly); **administrative savings assumptions**, while frequently the most politically visible point of contention, explain comparatively little of the variance in bottom-line national-spending results (Mercatus's own concession on this point is telling, given the institution's general skepticism of single payer); **scope/benefit-design differences** (whether LTSS, dental, vision, and zero cost-sharing are included) explain much of the divergence between the most favorable (Galvani, PERI) and least favorable (Urban Institute, Mercatus) national-spending estimates; **methodology** (static versus dynamic modeling) explains much of the divergence in distributional/household-impact estimates specifically (Heritage versus CRFB); and **data limitations** — specifically, the near-total absence of rigorous transition-cost, provider-labor-supply-response, and technological-innovation-effect modeling across every study reviewed — represent a category of disagreement that has not yet been adequately joined by the existing literature at all, rather than a dispute where credible analysts have already staked out and defended competing positions.

Consistent with the instruction not to present competing estimates as equally credible merely because they disagree: this review's assessment is that the CBO's illustrative-scenario framework, precisely because it is transparent about which assumptions drive which results and is produced by an institution with no stake in the political outcome, is the most useful single reference point for understanding the actual range of defensible outcomes — not because CBO is infallible, but because its methodology makes the assumption-sensitivity explicit rather than burying it inside a single headline number, which is the central analytical problem with most of the competing estimates on both sides.

11. Evidence Hierarchy

Ranked from strongest to weakest basis for causal inference, with what each type can and cannot establish for the specific questions this review addresses:

1. Randomized controlled trials / true natural-lottery experiments. *Example: the Oregon Health Insurance Experiment*. **Can establish:** causal effects of insurance coverage on utilization, financial protection, and (with sufficient statistical power) health outcomes, free from confounding, because treatment assignment is random. **Cannot establish:** effects at a scale, population, or benefit design different from the one studied (Oregon's experiment involved Medicaid specifically, a two-year window, and cannot be assumed to generalize to a

comprehensive national single-payer program); nor can a single RCT, however well-powered for its own purposes, resolve every downstream question (Oregon was underpowered for mortality/biomarker effects specifically).

2. Natural experiments with strong quasi-random variation and appropriate statistical design. *Examples: Taiwan's 1995 National Health Insurance implementation (joinpoint regression separating amenable from non-amenable mortality); Maryland's all-payer hospital rate regulation (difference-in-differences against comparison states); hospital-merger studies using merger-distance instruments (Cooper et al.).* **Can establish:** reasonably strong causal inference when the design includes an appropriate comparison group or a plausible reason to believe the "treatment" (a policy change) was not driven by the same factors as the outcome. **Cannot establish:** the same level of confidence as a true RCT, since some possibility of confounding or reverse causation typically remains, and the specific quantitative estimates are tied to the particular institutional and historical context studied.

3. Longitudinal and quasi-experimental studies using comparison groups. *Examples: Sommers et al.'s Massachusetts mortality study; Medicaid-expansion mortality studies (Miller, Johnson, Wherry).* **Can establish:** suggestive, policy-relevant evidence of causal effects, particularly when multiple independent studies using different comparison groups and time periods reach similar conclusions. **Cannot establish:** the same robustness as method 1 or 2; specific findings (such as the Massachusetts 2.9% mortality-reduction estimate) can be, and in this case have been, subject to legitimate methodological critique regarding the robustness of the comparison-group design.

4. Cross-national comparisons with careful confounder control and mechanism decomposition. *Examples: Papanicolas, Woskie & Jha (JAMA 2018); Anderson et al.'s "prices, not volume" studies.* **Can establish:** credible, policy-relevant descriptive findings about what structural factors (prices, administrative costs, utilization patterns) differ across systems, especially when a decomposition approach is used rather than simple aggregate comparison. **Cannot establish:** clean causal attribution to the financing model specifically, since countries differ in innumerable ways beyond financing structure; strongest when triangulated across multiple independent studies using different data and methods, as is the case for the "prices, not volume" finding specifically.

5. Cross-national outcome rankings without confounder adjustment. *Examples: the Commonwealth Fund's Mirror, Mirror composite ranking; simple life-expectancy league tables.* **Can establish:** that meaningful differences in outcomes exist and roughly how large they are. **Cannot establish:** that those differences are caused by the financing model rather than by demographic, behavioral, or social factors that also differ across the countries compared — a limitation this review's own Section 7 analysis addresses directly, and one the National Academies' 2013 panel explicitly cautions against glossing over.

6. Microsimulation models calibrated to real survey/claims data. *Examples: RAND's New York Health Act and national Medicare for All estimates; Urban Institute's HIPSM-based cost estimates; CBO's coverage and cost models.* **Can establish:** internally consistent, transparent projections of a specific policy design's likely effects, given the model's assumptions; useful for comparing the relative effects of different design choices within a consistent framework. **Cannot establish:** ground truth about untested real-world behavioral responses (how patients, providers, and employers would actually respond to a change of this scale and novelty in the U.S. context), since no microsimulation model can be validated against a directly comparable historical U.S. episode.

7. Actuarial/budgetary cost models. *Examples: Mercatus/Blahous; CMS NHE projections.* **Can establish:** disciplined, internally consistent dollar estimates given a specified set of assumptions, particularly useful for federal-budget-impact questions where the relevant actuarial methods are well-established. **Cannot establish:** behavioral or macroeconomic feedback effects unless explicitly and separately modeled (most actuarial models are more static than the general-equilibrium approach CBO's Working Paper 2022-02 uses).

8. Expert consensus panel reports. *Example: the National Academies' 2013 "Shorter Lives, Poorer Health."* **Can establish:** a rigorous, transparent synthesis of the existing literature and an authoritative judgment about which

questions remain genuinely unresolved — arguably most valuable precisely for identifying the boundaries of confident knowledge rather than for generating new primary findings. **Cannot establish:** new causal evidence beyond what the underlying literature it synthesizes already supports.

9. Systematic reviews of heterogeneous economic models. *Example: Cai et al.'s PLOS Medicine review of 22 single-payer cost studies*. **Can establish:** the central tendency and range of conclusions within the *published* literature on a topic. **Cannot establish:** which of the underlying, mutually inconsistent assumption sets is empirically more accurate — a systematic review of models that disagree because of assumptions, not data, inherits and can even obscure that underlying disagreement behind a headline percentage (e.g., "86% show savings") that describes the composition of published studies, not an independently validated empirical fact.

10. Advocacy-group analyses (whether pro- or anti-single-payer). *Examples: PNHP's cost and administrative-savings analyses; Heritage Foundation's distributional analysis; the Partnership for America's Health Care Future's messaging materials*. **Can establish:** a coherent, internally consistent argument for a predetermined position, often built on real underlying data; frequently useful for identifying which assumptions a given side considers most favorable to its case, which is itself informative. **Cannot establish:** a reliable, unbiased estimate of a contested empirical question, given the incentive to select assumptions favorable to the sponsoring organization's mission — this does not mean such analyses are wrong, but it means their conclusions should not be accepted without independent corroboration from a source without the same institutional stake, a standard this review has applied throughout (for example, treating the Himmelstein/Woolhandler qualitative finding as reliable specifically because it is independently corroborated by Papanicolas et al., a non-PNHP-affiliated research team, even while treating the specific magnitude of the PNHP-affiliated estimate with more caution).

11. Opinion pieces and single-outlet commentary. *Examples: think-tank blog posts and op-eds not accompanied by an underlying quantitative model*. **Can establish:** which arguments and framings are circulating in the policy debate and how specific evidence is being interpreted or contested by informed commentators (useful, for instance, for identifying the specific critiques of the Galvani/Yale study discussed in Section 5). **Cannot establish:** new evidence on its own; this review has used such sources only to identify critiques of, or context around, an underlying primary study, never as freestanding evidence for a quantitative claim.

This hierarchy is not a simple ranking of "trustworthy" versus "untrustworthy" institutions — several sources cited in the annotated bibliography that carry explicit institutional or advocacy affiliations (categories 9 and 10 above) nonetheless report real, independently verifiable underlying data, and several sources with no advocacy affiliation (category 6, microsimulation models) still carry irreducible modeling uncertainty. The hierarchy instead reflects what each type of evidence can and cannot establish about causation and about the likely real-world effects of a policy that has never been implemented at the scale and in the specific institutional context of the United States — which is the fundamental epistemic challenge underlying this entire review: no natural experiment or randomized trial can ever directly test "what would happen if the United States adopted single-payer financing nationally," and every source in this review is, at best, an informative but imperfect proxy for that unanswerable-by-direct-observation question.

12. What Does the Evidence Mean for the United States?

What is strongly supported by evidence?

The United States spends far more than any peer nation on health care because of higher prices and higher administrative overhead, not higher utilization. It underperforms peer nations badly on access, equity, and the specific mortality measures most directly tied to timely health-care access, while matching or outperforming them on several measures of clinical quality once care is actually delivered. Nearly half of lower- and middle-income Americans forgo needed care because of cost. Meaningful administrative savings under single-payer financing are essentially undisputed in direction, even among analysts otherwise skeptical of single payer. Comprehensive, benefit-rich single-payer designs of the kind most Medicare for All legislation has proposed

would substantially increase federal spending, on the order of $30 trillion or more over a decade — a finding two ideologically opposed institutions independently converge on.

What is probable but uncertain?

Whether total national health spending would rise or fall under a specific single-payer design is probable-but-uncertain in both directions, hinging on provider payment rates that are themselves a political and economic variable, not a fixed fact. Health outcomes on access-sensitive measures would probably improve; the effect on overall life expectancy is probably smaller than advocates suggest, given how much of the U.S. mortality disadvantage traces to factors outside health-care financing altogether. Some administrative savings are highly likely; the full magnitude often claimed by single-payer advocates is less certain. Employers would probably see reduced direct costs, probably passed through to wages over time, though the speed and completeness of that pass-through is uncertain.

What remains unknown?

Transition costs and risks at national scale are essentially unmodeled by the existing quantitative literature. How U.S. health-care providers — physicians facing large relative payment cuts especially — would actually respond in terms of labor supply, retirement timing, and practice patterns is not well studied. Whether existing provider capacity could absorb induced demand from universal, cost-sharing-free coverage without new access friction is genuinely unresolved. The long-run effect on pharmaceutical and medical-technology innovation is a legitimate, unresolved question on which this review found no rigorous, current, non-advocacy-affiliated quantitative answer.

The strongest arguments against single payer

The strongest evidence-based argument against single payer is not that it would fail to produce administrative savings or that other countries' systems don't work — both claims are contradicted by substantial evidence. It is that the transition risk is large and poorly quantified, that achieving net national savings requires provider payment reductions whose economic and political sustainability is genuinely uncertain (Vermont's abandoned 2011 plan is the most direct, if small-scale, cautionary evidence available), that the federal fiscal footprint of a comprehensive design is enormous by any accounting, and that supply-side capacity constraints could produce real access friction for a meaningful transition period even if long-run outcomes improved.

The strongest arguments for single payer

The strongest evidence-based argument for single payer is that it directly targets the two best-documented, most robustly replicated drivers of the U.S. cost and access crisis — fragmented, high-overhead multipayer administration and a health-care access gap concentrated among lower- and middle-income Americans — with a mechanism (universal, first-dollar public coverage) that the causal and quasi-experimental evidence reviewed here (Oregon, Taiwan, Massachusetts) associates with real improvements in financial protection, mental health, and amenable mortality. It would also resolve the specific structural problem, illustrated by Vermont's and New York's state-level struggles, that no individual U.S. state can achieve full single-payer efficiency without a national risk pool and federal cooperation.

Which criticisms are supported by evidence, and which are primarily assumption-driven?

Supported: concerns about transition risk and provider-payment-rate sustainability; concerns about the enormous scale of required federal financing; concerns that some of the most favorable savings and mortality claims (particularly the Galvani/Yale estimate) rest on assumptions independent fact-checkers have identified as unrealistic. Primarily assumption-driven rather than evidence-supported: the claim that administrative savings

are illusory or trivial (contradicted by independent, non-advocacy-affiliated research); the claim that other countries' lower costs prove nothing about price regulation's potential effectiveness in a U.S. multipayer context (contradicted by Maryland's own domestic experience); the claim that the U.S. health-care delivery system is simply worse than peer systems (contradicted by the U.S.'s comparably strong performance on several clinical-quality measures); and static distributional analyses that omit standard, well-established behavioral responses like wage pass-through.

What would have to be true for single payer to fail?

Provider payment cuts sufficient to generate meaningful savings would have to trigger significant reductions in provider supply or willingness to participate; induced demand from expanded, cost-sharing-free coverage would have to substantially outpace the health-care workforce's ability to expand or reallocate capacity; the transition itself would have to prove more disruptive and costly than any of the existing steady-state models anticipate, echoing Vermont's experience at a scale where recovery is much harder; and the financing burden would have to fall in a way that erodes political support before the system's benefits — many of which, per the international and natural-experiment evidence, take years to fully materialize — become visible to the public.

What would have to be true for it to succeed?

Administrative simplification would have to deliver savings closer to the higher end of the credible range (broadly consistent with, even if not fully matching, Canadian efficiency) without an offsetting collapse in provider participation; enough of the provider-payment reduction would have to be absorbed through reduced administrative burden and, for many providers, elimination of uncompensated care, rather than falling entirely on take-home provider income; the transition would have to be executed with enough operational competence and adequate lead time to avoid the kind of shock that undermined confidence in Vermont's plan; and the financing package would have to be, and be perceived to be, at least as fair as CRFB's more progressive financing-option findings suggest is achievable, so that the households bearing new costs are predominantly those most able to bear them.

Bottom-line conclusion

The evidence reviewed here does not support a simple yes-or-no answer to whether the United States could achieve lower total health-care costs and equal or better health outcomes through single-payer financing, and any claim that it does — from either direction — is overstating what the literature actually establishes. What the evidence supports is a more structured and, ultimately, more useful conclusion.

On cost, the United States spends dramatically more than any peer nation, and the evidence is unusually strong and consistent that this gap is driven by prices and administrative complexity rather than by higher utilization of care. This is among the most robust findings in all of health economics, replicated independently across multiple research teams and three decades. It follows directly from this finding that any reform capable of durably lowering U.S. health spending will need to address prices and administrative fragmentation specifically, not simply expand insurance coverage while leaving the underlying price and payment structure untouched. Single-payer financing is one credible mechanism for doing this, because a single national payer has the negotiating leverage and administrative uniformity that a fragmented multipayer system structurally lacks. But it is not the only mechanism: Maryland's all-payer hospital rate-setting system, Germany's and the Netherlands' all-payer fee schedules, and the federal government's own recent, limited exercise of Medicare drug-price negotiation authority all demonstrate that meaningful price and administrative reform is achievable within a multipayer structure, without the scale of disruption a full national single-payer transition would entail. This matters enormously for how the single-payer question should be framed: the choice is not simply "single payer versus the unreformed status quo," but rather a choice among several distinct paths to capturing the same underlying cost-saving mechanisms, each with a different risk, disruption, and political-feasibility profile.

On whether single payer specifically would reduce total national spending, the honest answer is that credible analysts, using transparent and defensible — if different — assumptions, reach genuinely different conclusions, and the disagreement traces almost entirely to two variables: how much provider payment rates would be cut, and how much administrative savings would actually be realized once transition costs and U.S.-specific structural frictions are accounted for. Neither variable is a fact waiting to be discovered by better research; both are policy choices that would be made, and contested, during actual implementation. This means that headline claims of "trillions saved" or "trillions in new costs" typically say more about which assumptions the source has chosen to feature than about a settled empirical reality. The Congressional Budget Office's own range — from roughly $740 billion in savings to $290 billion in additional cost over a single illustrative year, depending only on design choices within a single institution's own model — is the most honest single representation of this uncertainty available in the current literature.

On outcomes, the evidence supports real optimism about specific, well-defined benefits — reduced financial hardship, improved mental health, likely reductions in maternal and amenable mortality, and closing of the stark, well-documented access gap that leaves nearly half of lower-income Americans skipping needed care — because these are precisely the outcomes for which the causal and quasi-experimental evidence (Oregon, Taiwan, Massachusetts) is strongest and most directly tied to coverage and affordability. The evidence supports much more caution about claims that single payer would produce large, rapid gains in overall U.S. life expectancy or aggregate mortality, because a substantial share of the U.S. shortfall relative to peer nations is rooted in behavioral, social, and demographic factors that a financing-system change would not directly address, a conclusion the National Academies' own expert panel was explicit and careful about a decade ago and which remains the appropriate standard of caution today.

On implementation, this review's clearest independent finding is that the literature's near-total silence on transition costs, provider-labor-supply responses, and the operational risk of a national payment-system changeover is not a minor gap — it is arguably the single largest source of genuine uncertainty in the entire debate, larger than any of the disputed cost-modeling assumptions that dominate political discussion. Vermont's 2014 experience, in which a state that wanted to implement single payer and had a supportive legislative environment nonetheless found its anticipated administrative savings failing to materialize even before implementation began, is the most direct cautionary evidence available, and it argues for far more attention to implementation design and sequencing than the existing cost-modeling literature has generally provided.

It is also worth being explicit about a distinction this review has maintained throughout and that the political debate routinely collapses: evidence about "single-payer financing" as an abstract mechanism, drawn from Canada, Taiwan, and the domestic natural experiments reviewed here, is not the same as evidence about any specific Medicare for All bill, which typically bundles single-payer financing together with a much more comprehensive benefit design (including long-term care, dental, vision, and hearing) and the near-total elimination of cost-sharing than most existing single-payer systems abroad actually provide. A narrower single-payer design — closer in generosity to Canada's own system, which itself retains private financing for drugs, dental, and vision and permits some provincial cost-sharing outside its core benefit — would plausibly be markedly less expensive on a federal-budget basis than the $30-trillion-plus estimates this review's most-corroborated sources converge on, while a maximally comprehensive design would plausibly cost more. Readers evaluating any specific legislative proposal should therefore treat the cost and outcome ranges in this review as bounded by the scope of the benefit package actually proposed, not as a fixed verdict on "single payer" as an undifferentiated concept. This scope-sensitivity is not a hedge added to avoid a conclusion; it is one of the review's more load-bearing findings, since it is precisely the dimension along which the most favorable and least favorable cost estimates in the literature diverge.

It follows that the single most useful question a policymaker or engaged citizen can ask about any specific single-payer proposal is not "does the evidence support single payer," which this review has found to be the wrong level of generality to admit a clean answer, but rather a small number of narrower, evidence-answerable questions: what provider payment rates does this specific proposal assume, and are they sustainable without

meaningfully reducing provider participation; what benefit package does it cover, and how does that compare to the benchmark systems the evidence is drawn from; how is it financed, and how does that financing compare in distributional terms to the premiums, deductibles, and out-of-pocket costs it would replace; and what transition sequencing and lead time does it provide, given that transition risk, not steady-state economics, is this review's single largest identified source of uncertainty. A proposal that answers these questions credibly is meaningfully more evidence-supported than one that does not, regardless of which side of the single-payer debate it falls on — and much of the noise in the political debate on both sides comes from arguing about single payer in the abstract rather than engaging these specific, empirically tractable design questions.

Taken as a whole, the most defensible reading of the evidence is neither that single payer is a straightforward fiscal and moral improvement over the status quo, nor that it is fiscally reckless and doomed to fail. It is that single-payer financing is a plausible, evidence-consistent mechanism for capturing real, well-documented savings and access improvements that the current U.S. system leaves on the table, but that realizing those benefits at the national scale required depends on execution choices — provider payment design, transition sequencing, and financing structure — that the existing literature has not resolved and, in some respects, has barely begun to examine, and that the current U.S. multipayer system retains narrower, less disruptive paths to capturing some of the same underlying savings mechanisms should the political system prove unable or unwilling to attempt the larger transition.