The Invisible Wealth: Mismeasurement of Quality, the Myth of Stagnation, and the Underestimation of Real Income Growth

Franco Cazzaniga

Università dell'Insubria

franco.cazzaniga@uninsubria.it

December 12, 2025

Abstract

The dominant narrative of contemporary political economy holds that advanced economies have experienced decades of stagnating real wages and rising inequality, with the gains from technological progress accruing to a shrinking elite. This paper argues that this narrative rests on a statistical measurement illusion. Conventional national-accounts and consumer-price methodologies — designed for a mid-twentieth-century economy of standardized physical goods — systematically fail to capture the value generated by quality improvements, dematerialization, and the creation of entirely new goods and services.

We develop an attribute-based framework that reconceptualizes economic output as a flow of services from evolving bundles of characteristics rather than as quantities of nominally identical goods. Applying this framework to housing, durables, digital goods, and healthcare, we show that real living standards have improved far more rapidly than official statistics suggest. The consumer surplus from new and improved goods is economically equivalent to an expansion of real consumption possibilities: a worker with access to free navigation, global communication, and unlimited information commands a greater set of feasible choices than one without, even if their nominal incomes are identical.

The paper challenges the inevitability of Baumol's Cost Disease through a comparative analysis of two U.S. hospitals and a high-volume Indian tertiary-care center, showing that when institutional frictions are minimal, medical technology achieves substantially lower costs than prevailing U.S. benchmarks. The findings suggest that much of the "inequality crisis" reflects a crisis of measurement rather than a decline in real living standards.

Keywords: Real Wealth, Measurement Bias, Inequality, Hedonic Pricing, Dematerialization, Living Standards.

1 Introduction

The specter of inequality haunts contemporary economics. From Piketty's *Capital in the Twenty-First Century* (Piketty, 2014) to the distributional national accounts of Saez and Zucman (Saez and Zucman, 2016, 2020), a consistent narrative has emerged: the gains from economic growth have been captured by a narrow elite, real wages have stagnated for decades, and advanced societies are fracturing into a world of haves and have-nots. This narrative carries unmistakable echoes of nineteenth-century political economy—the immiseration thesis, the falling rate of profit, the inexorable concentration of capital.

This paper argues that the narrative is built on sand. Not because inequality of power and influence is illusory—it is not—but because the statistical apparatus used to measure material living standards is fundamentally broken. The System of National Accounts (SNA) and Consumer Price Index (CPI) methodologies were designed for a Fordist economy of standardized physical goods: identical automobiles rolling off assembly lines, identical refrigerators in identical kitchens. They are structurally incapable of capturing the value generated by an economy in which the goods themselves undergo continuous qualitative transformation, in which new categories of consumption emerge from nothing, and in which services once reserved for the wealthy become universally accessible at zero marginal cost.

1.1 The Invisible Wealth

Consider what a median worker in an advanced economy possesses today that was unavailable to the wealthiest person alive in 1970:

- instantaneous communication with anyone on Earth, at zero marginal cost;
- access to a body of knowledge exceeding any library in history;
- navigation capabilities superior to any military technology of the era;
- medical diagnostics (wearable sensors, imaging) once confined to hospitals;
- entertainment options (streaming, gaming) beyond any 1970 imagination;
- photographic capabilities exceeding professional equipment of the era.

None of this wealth appears in the wage statistics. The worker's "real income," as computed by deflating nominal wages by the CPI, may show stagnation or modest growth. But real income so computed is a fiction: it measures purchasing power over a basket of goods that no longer represents actual consumption, using price deflators that fail to account for the transformation of products and the emergence of entirely new categories.

The central thesis of this paper is that *real wealth has grown far more rapidly than real income as officially measured*. The gap between the two is not a rounding error; it is a chasm—potentially exceeding one percentage point of annual growth over several decades. If this estimate is even approximately correct, the narrative of secular stagnation and immiseration is not merely exaggerated; it is inverted. We are not poorer than we think. We are failing to count our riches.

1.2 The Inequality Illusion

The implications for the inequality debate are profound. The standard indictment—exemplified by Piketty's $r > g$ thesis and its derivatives (Piketty, 2014; Zucman, 2019)—relies on comparing income and wealth distributions over time, typically showing that gains have accrued disproportionately to the top deciles. But this comparison is only meaningful if the unit of measurement—"real income"—is stable across time. If a dollar of real income in 2025 purchases access to goods and services that were literally unavailable in 1975, then comparisons of real income distributions across these dates are comparing incommensurable quantities.

More concretely: the utility derived from the latest smartphone is identical for a billionaire and a median worker. Both have access to the same Google Maps, the same Wikipedia, the same streaming libraries. The billionaire cannot purchase a "better" Google—only more houses, more vehicles, more positional goods. The technology curve has *democratized* access to quality in precisely those domains—information, communication, entertainment, safety—where quality improvements have been most dramatic.

This observation does not deny that inequality exists. It reframes what inequality *means*. The gap in asset ownership (stocks, real estate, rare positional goods) has indeed widened. But assets confer two distinct benefits: consumption services and *power*—the capacity to influence decisions, shape institutions, and direct resources. The first has been substantially democratized by technology. The second has not.

The real divide in advanced economies is not between the wealthy and the impoverished, but between those who make institutional and corporate decisions and those who influence them through the market. The market is not an abstract exchange mechanism but a concrete social structure capable of exerting political force even on élites. It is the means through which dispersed economic actors impose constraints—price signals, budget limitations, and competitive pressures—that collectively discipline those who hold discretionary decision-making authority.

1.3 Outline of the Paper

The remainder of the paper proceeds as follows.

Section 2 confronts the "mismeasurement skeptics"—notably Syverson (2017) and Byrne et al. (2016)—who argue that measurement errors cannot explain the productivity slowdown. We show that their critique rests on an unwarranted equivalence between market output and economic welfare, and argue that their sectoral tests are mis-specified.

Section 3 develops the theoretical framework: an attribute-based model of economic output that treats goods as bundles of characteristics evolving through time. This framework provides the conceptual foundation for measuring the gap between official statistics and true living standards.

Section 4 applies these models to specific sectors—housing, appliances, food safety—where the disparity between measured prices and experienced utility is most pronounced.

Section 5 develops the formal decomposition of measurement bias into quality, dematerialization, and risk-adjustment components.

Section 6 draws out the implications for the inequality debate and for economic policy.

Section 7 concludes by revisiting the central paradox: we are measuring our wealth with instruments designed for a vanished economy, and the resulting statistics have become the foundation for a narrative of decline that bears little resemblance to lived experience.

Finally, the Appendices provide supplementary material.

Appendix A develops the formal foundations of the attribute-based framework, including intertemporal comparability, the multiplicative structure of Ω_t, aggregation properties, and the shadow-pricing interpretation of the effective budget constraint.

Appendix B applies the Institutional Friction Model to the comparative costing study of Erhun et al. (2020), showing that, once technological and regional price differences are accounted for, residual cost differentials in healthcare are primarily institutional rather than technological.

Appendix C provides an empirical illustration of the adjustment components, applying the framework to refrigerator energy efficiency, automobile safety, and smartphone dematerialization. The exercise demonstrates that each component of Ω_t is measurable using publicly available data and that the implied corrections are quantitatively large.

Appendix D outlines the empirical strategy, summarizing the data sources, measurement procedures, and identification considerations required to operationalize the framework.

2 The Mismeasurement Critique

A prominent strand of recent literature has challenged the view that mismeasurement can explain the observed productivity slowdown. Syverson (2017) argues that if unmeasured digital gains were substantial, we should observe faster productivity growth in IT-intensive sectors relative to others—yet no such differential appears in the data. Byrne et al. (2016) conduct a comprehensive accounting exercise and conclude that "while we find considerable evidence of mismeasurement, we find no evidence that, on balance, the understatement of

IT's contribution to productivity growth has gotten worse since the early 2000s."

This section argues that these critiques, while methodologically rigorous within their own framework, rest on a conceptual category error: the belief that "productivity" is synonymous with "measured output per hour in the market sector." Once this identification is questioned, the force of the critique dissolves.

2.1 The Sectoral Test and Its Hidden Assumption

2.1.1 A Formal Illustration of the Sectoral Mis-Specification

Syverson's sectoral test implicitly assumes that mismeasurement is proportional to the intensity of IT inputs used in production. Let $(A_{j,t})$ denote the time-indexed vector of attributes associated with good j. $A_{j,t}$ is the scalar index measuring the service-flow value of these attributes, and represents the economic contribution of the underlying characteristics (such as durability, efficiency, or safety). Let $(Q_{j,t})$ denote the vector of physical quantities or utilization measures of good j. With this convention, the effective service flow from good j at time t is

$$Y_{j,t} = Q_{j,t} \cdot A_{j,t}, \tag{1}$$

Measured productivity growth in sector j is:

$$g_{j,t}^{\text{meas}} = \frac{\dot{Y}_{j,t}}{Y_{j,t}} = \frac{\dot{Q}_{j,t}}{Q_{j,t}} + \frac{\dot{A}_{j,t}}{A_{j,t}}. \tag{2}$$

Syverson's test evaluates differences in $\dot{Y}_{j,t}/Y_{j,t}$ across sectors as a function of IT-intensity $I_{j,t}$, assuming that unmeasured productivity resides only in the production process:

$$g_{j,t}^{\text{meas}} = \alpha + \beta I_{j,t} + \varepsilon_{j,t}. \tag{3}$$

but mismeasurement arises from the evolution of attributes $A_{j,t}$, not from the intensity of IT inputs. The standard sectoral test evaluates whether sectors with greater IT intensity exhibit faster measured productivity growth. Our argument does not depend on claiming that quality improvements are orthogonal to IT inputs. In many cases, IT is precisely the instrument through which new attributes are engineered: safer airbags, more efficient motors, more reliable components. The point is not that IT and quality are unrelated, but that the *value of the attribute* is orthogonal to IT *productivity measurement*. Even when IT enables the improvement, the resulting attribute—safety, durability, efficiency—is not captured in the IT capital shares used by the sectoral test. The test asks whether IT-intensive sectors exhibit faster measured output growth, but the attribute-based framework shows that the relevant output is the flow of services generated by evolving characteristics, which remain invisible to the test.

If attribute growth $\dot{A}_{j,t}/A_{j,t}$ is orthogonal to IT-intensity—as is the case for improvements driven by materials science, chemistry, safety regulation, or design innovation—then:

$$\mathrm{Cov}\left(\frac{\dot{A}_{j,t}}{A_{j,t}},\ I_{j,t}\right) = 0, \tag{4}$$

and the sectoral regression is mechanically incapable of detecting the bias. The zero-covariance case illustrates the limiting scenario in which the test becomes maximally uninformative. More generally, mismeasurement remains undetectable whenever attribute evolution is weakly correlated with IT-intensity, not only when the correlation is exactly zero.

In particular, suppose the true productivity growth is:

$$g^*_{j,t} = \frac{\dot{Q}_{j,t}}{Q_{j,t}} + \frac{\dot{\Omega}_{j,t}}{\Omega_{j,t}}, \tag{5}$$

but the measured statistic omits $\Omega_{j,t}$. Then:

$$g^{\text{meas}}_{j,t} = g^*_{j,t} - \frac{\dot{\Omega}_{j,t}}{\Omega_{j,t}}, \tag{6}$$

and a sectoral test based on IT-intensity will find no differential even if $\dot{\Omega}_{j,t}/\Omega_{j,t}$ is large and heterogeneous.

The sectoral test, therefore, is mis-specified: it searches for mismeasurement in the production process rather than in the transformation of product attributes—the true locus of the bias.

Syverson's sectoral test presupposes that mismeasurement is a function of IT-intensity in the *production process*. If digital technology generates unmeasured value, the argument runs, this value should appear disproportionately in sectors that use IT intensively as an input.

But the framework developed in Section 2 implies a different locus of mismeasurement: the transformation of the *product itself*, not merely the process by which it is produced. An automobile manufactured in 2025 is safer than one manufactured in 1995 due to advances in metallurgy, structural engineering, airbag chemistry, and regulatory standards—not primarily due to onboard computers. Food is safer due to cold-chain logistics, HACCP protocols, and packaging materials. Housing is more energy-efficient due to insulation technology and window coatings.

These quality improvements are *orthogonal* to the IT-intensity of the production process. They arise from materials science, mechanical engineering, chemistry, and regulatory innovation. The sectoral test is therefore mis-specified: it looks for mismeasurement where IT is an input, when in fact mismeasurement occurs wherever the attribute vector of the output has improved faster than price deflators acknowledge.

2.2 The Market Boundary Fallacy

Byrne et al. (2016) explicitly acknowledge the welfare gains from digital services:

> For consumers, though, these services clearly enhance non-market leisure time and home production. This is important and valuable—our lives extend far beyond the borders of the market. Nevertheless, increases in household productivity for non-market services have never been counted in GDP and do not represent improved business productivity.

This passage reveals the conceptual category error at the heart of the critique. The authors correctly observe that free digital services enhance welfare, then dismiss this observation on the grounds that such gains "have never been counted in GDP." But the question at issue is precisely whether GDP *should* be interpreted as a measure of welfare or living standards—and if so, whether its current construction is adequate to that purpose.

Consider a worker earning $50,000 in 1990 and another earning $50,000 (in 1990 dollars) in 2025. The second worker has access to:

- instantaneous global communication at zero marginal cost;
- a navigation system superior to any 1990 technology at any price;
- an encyclopedia exceeding the Britannica in scope and accessibility;
- a high-resolution camera with unlimited exposures;
- a music library exceeding any 1990 collection;
- real-time translation services;
- access to the majority of human published knowledge.

The claim that these workers enjoy equivalent living standards—because their *measured* real incomes are identical—is not a methodological position but a definitional stipulation. It amounts to declaring, by fiat, that goods which escape market valuation do not contribute to economic welfare. This is the hallucination: treating an accounting convention as if it were a fact about the world.

2.3 Productivity of What?

The deeper issue concerns the purpose of productivity measurement. If the goal is to track the efficiency of market production narrowly construed, then Byrne, Fernald, and

Reinsdorf are correct: free digital goods are not "business productivity." But if the goal is to measure the economy's capacity to generate human welfare from given inputs of labor and capital, then the market boundary is arbitrary.

When Google Maps replaces a \$300 GPS device, GDP records a *contraction*—the GPS industry shrinks—while the consumer receives equivalent or superior functionality at zero cost. The productivity of the *system* has increased: fewer resources are required to deliver the same navigational service. That this gain manifests as "consumer surplus" rather than "business output" is an artifact of the pricing model (advertising-supported), not a fact about the underlying economic transformation.

More formally, let W_t denote aggregate welfare and Y_t measured GDP. The critique implicitly assumes:

$$\frac{\partial W_t}{\partial Y_t} = 1, \qquad \frac{\partial W_t}{\partial S_t} = 0, \tag{7}$$

where S_t denotes consumer surplus from non-market goods. The first condition states that welfare tracks measured output one-for-one; the second states that unmeasured surplus contributes nothing to welfare. Both are obviously false as descriptions of reality, yet both are required to sustain the conclusion that mismeasurement is economically irrelevant.

2.4 The Equivalence of Surplus and Income

The consumer surplus generated by free or quality-improved goods is economically equivalent to an increase in real income. If a service that previously cost \$1,000 per year is now available for free, the consumer's budget constraint has expanded by \$1,000—exactly as if they had received a \$1,000 raise.

The critique advanced here is not that GDP is an accounting identity—indeed, as an identity, it functions exactly as intended. The point is that the identity is often treated as a proxy for value creation or economic development. When a safer car or more durable appliance is sold at the same nominal price as its predecessor, measured output is flat, but value creation has increased. GDP is an adequate measure of market transactions; it is a poor measure of the value generated by qualitative improvements, new goods, and risk reductions. The problem is not with the identity itself, but with its interpretation as a welfare or development indicator.

Let p_0 be the historical price of a service and $p_1 = 0$ its current price (or, more generally, $p_1 < p_0$ after quality adjustment). The welfare gain is:

$$\Delta W = \int_{p_1}^{p_0} x(p)\, dp + (p_0 - p_1) \cdot x(p_1), \tag{8}$$

where $x(p)$ is the demand function. The first term is the Hicksian consumer surplus; the second is the direct budget savings. For a good that becomes free, this reduces to:

$$\Delta W = \int_0^{p_0} x(p)\, dp, \tag{9}$$

which can be substantial—potentially exceeding the original expenditure—if demand is elastic or if the good was previously unaffordable.

It is important to emphasize that consumer surplus does not constitute monetary income: a household cannot use Google Maps or Wikipedia to pay rent. However, this observation misstates the relevant economic mechanism. Intuitively, digital services collapse search costs, reduce uncertainty, and eliminate many of the frictions that previously inflated the effective cost of consuming complementary goods. As a result, the shadow price—the marginal effective cost faced by the consumer— declines even when nominal prices remain unchanged. Consumer surplus lowers the *shadow price* of many complementary goods and activities by collapsing search costs, reducing information frictions, minimizing waste, and enabling local optimization of consumption choices. A consumer cannot convert digital services into cash, but can use them to access cheaper stores, avoid costly mistakes, substitute away from inferior products, or improve the efficiency of time allocation.

In classical welfare analysis, these adjustments are invisible because the budget constraint is assumed to take the form $p \cdot q \leq y$, with perfect monetary substitutability between all goods. In reality, the household budget constraint is *locally* determined by transaction costs, information costs, uncertainty, and the availability of substitute activities. Digital goods reduce these frictions dramatically. Hence, while consumer surplus is not income in the accounting sense, it *modifies the effective budget constraint* by lowering the shadow cost of achieving a given level of consumption. The neoclassical identification of "real income" with deflated nominal expenditure obscures this mechanism and leads to a systematic understatement of welfare gains.

This surplus is *real income* in every economically meaningful sense. The fact that it does not appear in National Accounts is a measurement failure, not evidence that the gains are illusory. A formal treatment of this mechanism, based on shadow prices and the effective budget constraint, is provided in Appendix B.5.

2.5 Why the Critique Misses the Point

The mismeasurement skeptics pose the question: "Has mismeasurement *increased* since 2004?" If not, they argue, mismeasurement cannot explain the productivity *slowdown*.

But this framing presupposes that measured productivity growth before 2004 was approximately correct. The argument of this paper is different: productivity has been *systematically undermeasured throughout the postwar period*, because quality improvements and dematerialization have consistently outpaced the adjustments embedded in official deflators.

The relevant question is not whether mismeasurement has accelerated, but whether it is large in absolute terms. If true productivity growth has been 1.5 percentage points per year higher than measured productivity growth for decades, then the narrative of "secular stagnation" is simply false—regardless of whether the bias has widened or narrowed.

2.6 A Paradigm Distinction

To summarize: the mismeasurement critique operates within a paradigm that identifies economic welfare with measured market output. Within this paradigm, the critique is internally consistent. But the paradigm itself is the hallucination.

The framework developed in this paper operates within a different paradigm: economic welfare depends on the *flow of services* derived from goods and activities, whether or not these services are priced in markets. From this perspective, the distinction between "business productivity" and "household productivity" is an accounting artifact, not an economic fundamental.

The question is not whether free digital goods "count" as productivity—that is a matter of definition. The question is whether our measures of living standards accurately capture the transformation in human welfare over the past half-century. The evidence presented in subsequent sections suggests they do not.

3 Theoretical Framework

The attribute-based framework developed in this section establishes that the adjustment factor Ω_t decomposes into quality, dematerialization, and risk-reduction components. This approach builds on the characteristics-based view of goods introduced by Lancaster (1966) and the hedonic valuation framework of Rosen (1974), extending these foundations to intertemporal welfare measurement.

This section operationalizes that framework by developing concrete measurement strategies for each component. We focus on two domains where the gap between measured and true value is most pronounced: durable goods (where quality improvements dominate) and digital convergence (where dematerialization dominates).

3.1 Quality-Adjusted Service Flow (Durables and Housing)

For durable goods, the relevant economic magnitude is not the purchase price but the *service flow*—the stream of utility derived from ownership over the asset's lifetime. Standard inflation metrics track the purchase price P_t, implicitly treating the good as unchanged across time. The attribute-based framework requires instead that we track the cost of obtaining a unit of service, adjusted for the evolution of the quality vector $\mathbf{q}_t$.

3.1.1 The User Cost of Capital

Following the standard theory of durable goods (Jorgenson, 1963), the user cost of capital for a durable asset is:

$$\rho_t^{\text{nom}} = P_t(r_t + \delta_t - \pi_t), \tag{10}$$

where r_t is the real interest rate, δ_t the depreciation rate, and π_t the expected rate of capital gain. This expression gives the annual cost of holding one unit of the asset.

However, if the asset's quality has improved—captured by the quality multiplier $\theta_t \equiv v_q(\mathbf{q}_t)/v_q(\mathbf{q}_0)$ defined formally later in Equation (19)—then the effective user cost per unit of *service* is:

$$\rho_t = \frac{P_t(r_t + \delta_t - \pi_t)}{\theta_t}. \tag{11}$$

The denominator θ_t reflects the fact that a higher-quality asset delivers more service per unit of ownership. A refrigerator that consumes 60% less energy, lasts twice as long, and preserves food more effectively is not "the same good" as its 1980 predecessor; it delivers more service per dollar of user cost.

3.1.2 The Quality Bias

Statistical agencies implicitly hold $\theta_t = 1$ for all t, which suppresses quality improvements from the price index. The resulting measurement bias can be derived by comparing the growth rates of nominal and quality-adjusted user costs.

Taking logarithms of Equation (11) and differentiating:

$$\frac{\dot{\rho}_t}{\rho_t} = \frac{\dot{P}_t}{P_t} + \frac{d}{dt}\ln(r_t + \delta_t - \pi_t) - \frac{\dot{\theta}_t}{\theta_t}. \tag{12}$$

If the financial terms $(r_t + \delta_t - \pi_t)$ are approximately stable, the quality bias reduces to:

$$B_{q,t} = \frac{\dot{P}_t}{P_t} - \frac{\dot{\rho}_t}{\rho_t} \approx \frac{\dot{\theta}_t}{\theta_t}. \tag{13}$$

This confirms the result stated in Section 2: official inflation is overstated, and real income growth understated, by the rate of quality improvement.

3.1.3 Empirical Considerations

Several empirical regularities amplify the quality bias beyond what Equation (13) suggests:

1. **Declining depreciation rates.** Improved engineering has extended the lifespan of major appliances and vehicles. If δ_t falls over time, the denominator $(r_t + \delta_t - \pi_t)$ shrinks, reducing nominal user cost even before quality adjustment.

2. **Operating cost reductions.** Energy efficiency improvements reduce the total cost of ownership but are not captured in purchase price indices. A quality-adjusted measure should incorporate lifetime energy savings.

3. **Safety improvements.** Reduced probability of injury or death constitutes a welfare gain that standard deflators ignore entirely. This is captured by the risk-reduction component which is defined so that $(1 + \rho_t)$ is the multiplicative adjustment factor.

The Boskin Commission (Boskin et al., 1996) estimated quality bias at approximately 0.6 percentage points per year in the mid-1990s. Subsequent work by Bils and Klenow (2001) and Pakes (2003) suggests the bias may be larger for categories experiencing rapid innovation. The framework developed here provides a unified theoretical foundation for these empirical estimates.

3.2 The Dematerialization Surplus (Digital Goods)

The second major source of mismeasurement arises when a single digital device replaces a basket of previously distinct physical goods. This is captured by the dematerialization component δ_t in the adjustment factor Ω_t.

3.2.1 The New Goods Problem

The challenge is a generalization of what Hausman (1996) termed the "new goods problem": how should we value a good that did not exist in the base period? The standard approach—introducing the new good at its first observed price—misses the entire consumer surplus generated by its creation.

Digital convergence intensifies this problem. A smartphone is not merely a "new good"; it is a *bundle of functionalities* that previously required separate physical devices: camera, GPS navigator, encyclopedia, audio system, communication device, and more. When the smartphone enters the market, the physical industries producing these devices contract or disappear. GDP records this as a decline in economic activity, even as consumer welfare has increased.

3.2.2 Formal Definition

Let $G_t = \{g_1, g_2, \ldots, g_K\}$ denote the set of physical goods historically required to provide the functionalities now bundled into a digital device n. Let P_{g_k,t_0} denote the price of good g_k at the base period t_0, and let $P_{n,t}$ denote the current price of the digital device.

The **Dematerialization Surplus** is defined as:

$$DS_t = \underbrace{\sum_{k=1}^{K} P_{g_k,t_0}}_{\text{displaced physical goods}} - \underbrace{P_{n,t}}_{\text{convergent device}} + \underbrace{\int_{P_{n,t}}^{\bar{p}} x_n(p)\, dp}_{\text{consumer surplus}} \tag{14}$$

where $x_n(p)$ is the demand function for the digital device and $\bar{p}$ is the choke price (the price at which demand falls to zero) [1] . The first two terms capture the direct cost savings from convergence: the consumer no longer needs to purchase K separate devices. The third term captures the additional surplus from functionalities that were previously unavailable at any price—real-time navigation, instant global communication, access to the world's information.

It is important to distinguish the two components of Equation Equation (14). The term $\sum_{k=1}^{K} P_{g_k,t_0} - P_{n,t}$ represents a *resource-release effect*: the consumer no longer needs to allocate expenditure to the displaced physical goods. The integral $\int_{P_{n,t}}^{\bar{p}} x_n(p)\, dp$ represents the *new utility effect* associated with functionalities that previously did not exist. This decomposition follows the logic of Hausman (1996): the total value of the innovation is the access to the new capability plus the resources freed for alternative uses. There is no double counting, because the integral covers only the valuation of the new attributes, not the benefit arising from no longer purchasing the displaced goods.

3.2.3 Relation to the Attribute Framework

The dematerialization surplus corresponds to the term $(1 + \delta_t)$ in the multiplicative decomposition of Ω_t. Specifically:

$$1 + \delta_t = \frac{v_d(\mathbf{d}_t)}{v_d(\mathbf{d}_0)}, \tag{15}$$

where $\mathbf{d}_t$ is the vector of dematerialized functionalities available at time t. At the base period, $\mathbf{d}_0$ may be null (no dematerialization) or minimal; as digital convergence proceeds, $\mathbf{d}_t$ expands and δ_t grows.

The key insight is that standard GDP calculations capture only $P_{n,t}$—the price of the convergent device. The displaced physical goods vanish from the accounts (as their industries contract), and the consumer surplus is never recorded. This produces the paradox noted in Section 1: an increase in consumer welfare is recorded as a contraction in measured output.

[1]The integral represents a Hicksian compensating variation and is therefore measured in monetary units. All three components of Equation (5) are expressed in comparable units, ensuring dimensional consistency.

3.2.4 Magnitude Estimates

To illustrate the potential magnitude, consider the smartphone. Brynjolfsson et al. (2019) estimate that the median consumer would require compensation exceeding \$10,000 per year to forgo access to search engines alone. Similar estimates for mapping services, social media, and messaging suggest that the unmeasured consumer surplus from digital goods may rival or exceed measured GDP growth in recent decades.

These estimates are necessarily imprecise, but they establish that the dematerialization surplus is not a rounding error. It represents a fundamental category of value that the System of National Accounts was not designed to capture.

3.3 Integrating the Components

The quality-adjusted user cost and dematerialization surplus are not independent corrections; they are applications of the unified framework developed in Section 2. For any good or service, the true effective price is:

$$P_t^\dagger = \frac{P_t}{\Omega_t} = \frac{P_t}{\theta_t \cdot (1+\delta_t) \cdot (1+\rho_t)}, \tag{16}$$

where each component captures a distinct source of unmeasured value.

3.4 Formal Definition of the Adjustment Factor Ω_t

The attribute-based framework defines observed prices as incomplete measures of the true economic cost of obtaining a flow of services. Let x_t denote a good at time t, represented as a vector of attributes:

$$x_t = \big(q_t, d_t, s_t\big), \tag{17}$$

where q_t is an index of quality-enhancing characteristics (durability, efficiency, safety), d_t captures dematerialized functionalities, and s_t represents the reduction in risk or hazard per unit of consumption.

We define the *effective price* of the good as:

$$P_t^\dagger = \frac{P_t}{\Omega_t}, \tag{18}$$

where P_t is the observed nominal price and Ω_t is the adjustment factor that translates nominal expenditures into units of service flow.

Formally, The adjustment factor Ω_t is decomposed multiplicatively as:

$$\Omega_t = \theta_t \cdot (1+\delta_t) \cdot (1+\rho_t), \tag{19}$$

where:

- θ_t measures quality improvements in the service flow per unit of expenditure;
- δ_t measures gains from dematerialization—the replacement of physical assets with digital functionalities;
- ρ_t measures reductions in the probability or severity of adverse outcomes.

The multiplicative form reflects the fact that improvements in distinct attributes increase the effective service flow independently. In the limit where $(\theta_t, \delta_t, \rho_t) = (1, 0, 0)$, the framework collapses to the standard treatment of goods as homogeneous and time-invariant, as implicitly assumed in the System of National Accounts.

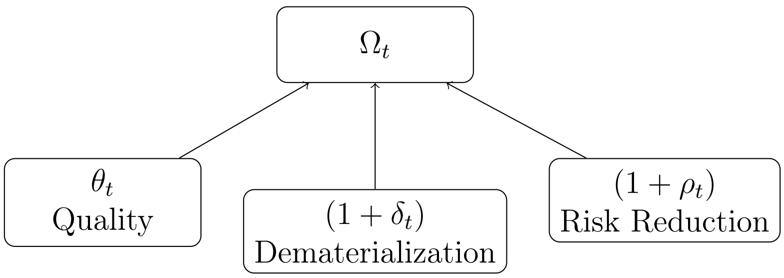

Figure 1: Decomposition of the adjustment factor Ω_t into quality, dematerialization, and risk-reduction components.

For traditional durables (appliances, vehicles, housing), the dominant adjustment is θ_t—quality improvements that increase service flow per unit of cost. For digital goods, the dominant adjustment is δ_t—the bundling of previously separate functionalities into a single device. For goods where safety has improved dramatically (food, transportation, medical devices), the risk-reduction term ρ_t becomes material.

The total measurement bias, aggregated across the economy, is the subject of Section 6. First, we examine specific sectors where these biases are most pronounced.

4 Sectoral Analysis: The Hidden Deflation

Having established the theoretical framework and measurement methodology, we now apply these tools to specific sectors where the disparity between measured inflation and experienced welfare is most pronounced. Each case illustrates a different component of the adjustment factor Ω_t: quality improvements in housing and durables (θ_t), and risk reduction in food safety $(1 + \rho_t)$.

4.1 Housing and Durables: The Efficiency Dividend

In the housing market, price indices typically control for square footage and location but struggle to account for systemic improvements in habitability standards. A residential unit constructed in 2024 is a fundamentally different asset from one built in 1970: thermal insulation, electrical safety, air quality systems, and structural resilience have all improved dramatically.

In terms of the attribute framework, the quality vector $\mathbf{q}_t$ for housing includes energy efficiency, safety compliance, durability, and maintenance requirements. Each of these has shifted favorably over time, implying $\theta_t > 1$ and growing.

4.1.1 Energy Efficiency

The most quantifiable improvement is energy consumption. Applying the quality-adjusted user cost model from Equation (11), we observe that while the nominal price P_t of housing has risen substantially, the energy input E_t required to maintain a comfortable temperature has collapsed. Modern insulation standards, double-glazed windows, and efficient HVAC systems have reduced heating and cooling costs per square meter by 40–60% relative to 1970s construction.

This reduction in operating cost is economically equivalent to a price decrease: the *total cost of ownership*—purchase price plus lifetime operating costs—has fallen relative to the service delivered. Yet standard price indices capture only the purchase price, missing the efficiency dividend entirely.

4.1.2 Household Appliances

For household appliances, the efficiency dividend is even more pronounced. Consider the refrigerator:

- A modern unit costs approximately the same in nominal terms as its 1980 equivalent, but a fraction of the cost in labor-hours.
- Energy consumption has fallen by approximately 75%, reducing lifetime operating costs by thousands of dollars (see Appendix C for details).
- Preservation technology has improved, reducing food waste and extending storage duration.
- Reliability has increased, with mean time between failures substantially longer than four decades ago.

If the CPI treats a "refrigerator" as a static good across time, it misses all of these improvements. The quality multiplier θ_t for refrigerators is substantially greater than unity, implying that the effective price—the cost per unit of refrigeration service—has fallen far more rapidly than the nominal price.

Similar patterns hold for washing machines, air conditioners, automobiles, and virtually every category of durable good. The cumulative effect is a systematic overstatement of inflation and understatement of real income growth.

4.2 Risk-Adjusted Consumption: The Case of Food Safety

A critical but often overlooked component of living standards is the reduction of risk. Standard metrics evaluate food prices based on caloric or weight units, treating "food" as a homogeneous commodity. But the consumer is not merely purchasing calories; they are purchasing *safe* calories.

4.2.1 The Risk-Adjustment Framework

In terms of the attribute framework, food safety is captured by the risk-reduction component $(1 + \rho_t)$. Let π_t denote the probability of foodborne illness per unit of consumption, and let C_{loss} denote the expected economic damage from such an event (medical costs, lost labor, mortality risk). The risk-adjusted price is:

$$P_t^{\text{adj}} = P_t^{\text{nom}} + \pi_t \cdot C_{\text{loss}}. \tag{20}$$

The second term represents the implicit "risk premium" embedded in food consumption—the expected cost of adverse outcomes.

4.2.2 The Collapse of Food Risk

In the post-war era, π_t was substantial. Foodborne illness was common; refrigeration was unreliable; supply chains were short and variable; contamination events were frequent. Today, thanks to cold chain logistics, HACCP protocols, pasteurization, and advanced packaging technology, $\pi_t \to 0$ for most food categories in developed economies.

This collapse of the risk premium represents a massive welfare gain that is invisible to standard price indices. The nominal price of food may have risen modestly, but the risk-adjusted price—the cost of obtaining safe nutrition—has fallen dramatically. A consumer in 2025 faces a probability of serious foodborne illness that is orders of magnitude lower than a consumer in 1955, yet this improvement appears nowhere in the statistics.

4.2.3 Generalization

The food safety example generalizes to other domains where risk has declined:

- **Automobile safety**: Fatality rates per mile driven have fallen by roughly 80% since 1970, due to structural improvements, airbags, seatbelt adoption, and crash-avoidance technology.

- **Pharmaceutical safety**: Modern drug approval processes and manufacturing standards have reduced the incidence of adverse drug reactions.
- **Consumer product safety**: Electrical safety standards, flame-retardant materials, and child-safety requirements have reduced household accidents.

In each case, the reduction in π_t constitutes a welfare gain that standard deflators fail to capture. The consumer is not merely buying the same good at a different price; they are buying a *safer* good, and the safety improvement has economic value.

4.3 Summary

The sectoral evidence confirms the theoretical framework: quality improvements (θ_t) and risk reductions (ρ_t) generate substantial welfare gains that are invisible to standard price indices. For housing and durables, the efficiency dividend implies that the true cost of living has fallen relative to official measures. For food and other risk-sensitive categories, the collapse of the risk premium implies a similar understatement of welfare growth.

These are not marginal corrections. Cumulated over decades, they imply that real living standards have improved far more rapidly than official statistics suggest—a finding with profound implications for the inequality debate.

5 Aggregate Mismeasurement: Magnitudes and Implications

The preceding sections established the theoretical framework and applied it to specific sectors. This section aggregates these components to estimate the total measurement bias in the macroeconomic accounts and examines its implications for our understanding of economic growth.

5.1 Aggregating the Bias Components

Recall from Section 2 that the total adjustment factor for any good is:

$$\Omega_t = \theta_t \cdot (1 + \delta_t) \cdot (1 + \rho_t), \tag{21}$$

where θ_t captures quality improvements, δ_t dematerialization gains, and ρ_t risk reduction. The aggregate measurement bias B_t depends on the distribution of these factors across the consumption basket and their rates of change over time.

Bias Component	Symbol	Source of Gain	Empirical Range (pp/yr)
Quality improvements	θ_t	Durability, efficiency, safety	0.4–0.8
Dematerialization	δ_t	Digital convergence, new goods	0.2–0.5
Risk reduction	ρ_t	Hazard collapse (food, transport)	0.1–0.3
Total bias	Ω_t	Multiplicative combination	0.8–1.5

Table 1: Summary of bias components and approximate magnitudes.

For a first-order approximation, the growth rate of the aggregate bias is:

$$\frac{\dot{B}_t}{Y_t} \approx \sum_i \omega_i \left(\frac{\dot{\theta}_{i,t}}{\theta_{i,t}} + \frac{\dot{\delta}_{i,t}}{1+\delta_{i,t}} + \frac{\dot{\rho}_{i,t}}{1+\rho_{i,t}} \right), \tag{22}$$

where ω_i is the expenditure share of good i and the sum runs over all consumption categories. This is a first-order approximation that abstracts from changes in expenditure shares. A full Divisia index would account for time-varying weights, but such refinements do not affect the qualitative result that unmeasured attribute growth increases real consumption possibilities.

5.2 Existing Estimates

Several independent lines of research provide estimates of the component biases:

5.2.1 Quality Bias

The Boskin Commission (Boskin et al., 1996) estimated that quality improvements contributed approximately 0.6 percentage points per year to upward bias in the CPI during the mid-1990s. Subsequent research has both challenged and extended this estimate:

- Bils and Klenow (2001) found evidence of substantial quality upgrading within product categories, suggesting the Boskin estimate may be conservative for goods experiencing rapid innovation.
- Pakes (2003) developed refined hedonic methods that capture quality change more accurately, generally confirming significant bias.
- Gordon (2016) argues that quality improvements have slowed in recent decades, though this assessment focuses primarily on IT goods and may miss improvements in safety, durability, and operating efficiency.

A reasonable central estimate for the quality component $\dot{\theta}/\theta$ is 0.4–0.8 percentage points per year, with substantial variation across product categories.

5.2.2 Dematerialization Bias

The dematerialization component is harder to estimate because it involves goods that have disappeared from the accounts entirely. Brynjolfsson et al. (2019) provide experimental estimates of consumer willingness-to-pay for digital services:

- Median willingness-to-accept for giving up search engines: \$17,530 per year.
- Median willingness-to-accept for giving up digital maps: \$3,648 per year.
- Median willingness-to-accept for giving up email: \$8,414 per year.

These figures suggest that unmeasured consumer surplus from digital goods alone may exceed several percentage points of GDP. However, translating willingness-to-pay into bias estimates requires assumptions about how this surplus would have been measured had the goods been priced conventionally.

A conservative estimate for the dematerialization component $\dot{\delta}/(1+\delta)$ is 0.2–0.5 percentage points per year, concentrated in the period since widespread internet adoption.

5.2.3 Risk-Reduction Bias

The risk-reduction component has received less systematic attention, but specific studies provide indicative magnitudes:

- Cutler and McClellan (2001) estimate that improvements in medical technology have generated welfare gains substantially exceeding measured healthcare expenditure growth.
- Automobile safety improvements (fatality rates down 80% since 1970) represent welfare gains on the order of hundreds of billions of dollars annually, using standard value-of-statistical-life calculations.
- Nordhaus (1996) demonstrates that the true cost of illumination fell by a factor of 1,000 over two centuries—a decline almost entirely missed by conventional price indices.

A rough estimate for the risk-reduction component $\dot{\rho}/(1+\rho)$ is 0.1–0.3 percentage points per year, though this is highly uncertain.

5.3 Total Bias Estimate

Summing the components yields a central estimate of aggregate measurement bias in the range of 0.8–1.5 percentage points per year. This estimate is broadly consistent with:

- The original Boskin Commission estimate of 1.1 percentage points (for the CPI, not GDP deflator).
- Nordhaus (1996), whose historical analysis of lighting prices implied enormous unmeasured welfare gains.
- Brynjolfsson et al. (2019), whose experimental methods suggest large unmeasured consumer surplus from digital goods.

It is inconsistent with the conclusions of Byrne et al. (2016) and Syverson (2017), who argue that mismeasurement cannot explain the productivity slowdown. However, as argued in Section 3, their critique addresses a different question: whether mismeasurement has *accelerated* since 2004. Our argument is that mismeasurement has been *large in absolute terms* throughout the period, implying that true living standards have grown faster than official statistics suggest—regardless of whether the bias has widened or narrowed.

An empirical illustration applying these components to specific goods is provided in Appendix C. Using publicly available data on refrigerator energy efficiency, automobile fatality rates, and smartphone functionality, the exercise demonstrates that each component of Ω_t is directly measurable and that the implied corrections are quantitatively large—on the order of a factor of 1.5 to 4.0 for individual goods. While the illustration does not constitute a comprehensive deflator, it confirms that the magnitudes suggested by the Boskin Commission and subsequent literature are empirically grounded rather than speculative.

5.4 Implications for Real Income Growth

If the aggregate bias is approximately 1 percentage point per year, the implications are profound:

1. **Cumulative effect**: Over 40 years, a 1 percentage point annual bias compounds to a level difference of approximately 50%. Real incomes in 2025 may be roughly 50% higher than official statistics suggest, relative to 1985.
2. **Wage stagnation narrative**: The claim that median real wages have stagnated since the 1970s relies on deflating nominal wages by the CPI. If the CPI overstates inflation by 1 percentage point per year, then real wages have in fact grown substantially—just not as recorded.

3. **Inequality dynamics**: Comparisons of real income distributions across decades are comparing incommensurable quantities. A "stagnant" real income in 2025 purchases access to goods, services, and safety levels that did not exist in 1985.

4. **International comparisons**: Countries with different deflation methodologies may show spurious differences in real income growth. The apparent "productivity miracle" in some countries and "stagnation" in others may partly reflect measurement conventions.

5.5 Caveats and Limitations

Several caveats apply to these estimates:

- **Heterogeneity**: The bias varies substantially across product categories and may not apply uniformly to all consumers. Those whose consumption is concentrated in categories with slow quality improvement (e.g., housing in supply-constrained markets) may experience less unmeasured welfare gain.
- **Substitution**: Some unmeasured gains may represent substitution rather than net welfare increase—time spent on smartphones may displace other activities with positive value.
- **Distribution**: Even if aggregate welfare has grown faster than measured, the distribution of unmeasured gains may differ from the distribution of measured income. Digital goods are relatively egalitarian (everyone gets the same Google), but quality improvements in housing may accrue disproportionately to owners of newer properties.

These caveats do not overturn the central finding—that measured real income substantially understates true welfare growth—but they caution against mechanical application of aggregate bias estimates to distributional questions.

6 Implications: Rethinking Inequality

The findings of this paper have profound implications for how we understand inequality, economic progress, and the relationship between the two. If real living standards have grown substantially faster than official statistics suggest, then the dominant narrative of stagnation and immiseration requires fundamental revision.

6.1 The Incommensurability Problem

The inequality literature relies on comparing income and wealth distributions across time. Piketty (2014) documents the share of income accruing to the top deciles; Saez and

Zucman (2016) traces wealth concentration over a century. These are valuable empirical contributions. But they rest on an implicit assumption: that "real income" and "real wealth" are comparable across decades.

The attribute-based framework developed in this paper challenges that assumption. If a dollar of real income in 2025 purchases access to goods, services, and safety levels that were unavailable at any price in 1975, then comparisons of real income distributions across these dates are comparing incommensurable quantities.

Consider a concrete example. A worker at the 50th percentile of the income distribution in 1975 could purchase:

- no portable communication device;
- no access to real-time navigation;
- no instantaneous access to recorded music beyond their personal collection;
- no ability to search the world's information;
- food with substantially higher contamination risk;
- an automobile with an order of magnitude higher fatality risk per mile.

A worker at the 50th percentile in 2025 possesses all of these capabilities, most at zero marginal cost. To claim that these two workers have "equivalent" living standards because their CPI-deflated incomes are similar is to privilege an accounting convention over economic reality.

6.2 The Democratization of Quality

One of the most striking features of modern technological progress is its egalitarian character with respect to quality. The smartphone in the pocket of a minimum-wage worker contains the same search engine, the same mapping software, the same communication capabilities as the smartphone owned by a billionaire. The antibiotics prescribed to a Medicaid patient are chemically identical to those prescribed to a Fortune 500 CEO. The safety features in a Honda Civic—airbags, antilock brakes, stability control—are functionally equivalent to those in a Mercedes. The claim concerns the quality of the service conditional on access, not the institutional process that governs access. What converges is the technological capability embedded in the good, not the structure of healthcare provision itself.

This democratization of quality is historically unprecedented. In prior eras, quality gradations were continuous: the wealthy ate better food, wore better clothes, traveled in greater comfort and safety, and had access to better medical care at every margin. Today, for a wide range of goods and services, quality has converged at the top. The rich can purchase *more*—more houses, more cars, more positional goods—but they cannot purchase *better* in many domains that matter for daily welfare.

The implications for inequality measurement are significant. Standard indices track the distribution of purchasing power over a fixed basket. But if the quality of goods available to the median consumer has converged toward the quality available to the wealthy, then inequality in *experienced welfare* has narrowed even as inequality in measured income has widened.

The argument does not require the measurement bias to be larger for lower-income households. What matters is that many of the gains ignored by official statistics—information, communication, navigation, safety, and basic digital services—exhibit strong quality convergence across the income distribution. Inequality in measured income may well rise while inequality in experienced quality declines, because the technologies that generate the unmeasured surplus are essentially non-excludable and non-rival at the point of use. The "inequality illusion" refers to the conflation of these two distinct dimensions.

The convergence of quality across the income distribution implies convergence in the experienced components of welfare, even when measured incomes diverge. This does not mean that all forms of inequality vanish; it means that income dispersion and welfare dispersion are no longer aligned. To understand what remains unequal, it is necessary to separate the material dimension of living standards from the institutional and political dimension of power. A qualification is in order. The democratization of quality documented above does not apply uniformly across all categories of consumption. A growing empirical literature shows that lower-income households face systematically higher inflation for basic consumer packaged goods, raising the question of how these distributional price dynamics interact with the attribute-based gains identified in this paper. Although recent literature shows that lower-income households face higher inflation for basic consumer packaged goods, this is a short-term phenomenon driven by the effect of inflationary shocks on different markets and does not seem to be related to the general trend studied here.

6.3 A Note on Inflation Inequality

Recent empirical work documents that lower-income households face systematically higher inflation for consumer packaged goods. Jaravel (2019) finds that annual inflation was approximately 0.6–0.9 percentage points higher for the bottom income quintile relative to the top quintile, driven by the direction of product innovation toward premium market segments. Separately, Cavallo and Kryvtsov (2024) document "cheapflation" during the post-pandemic inflation surge, with prices of budget product varieties rising faster than premium alternatives.

These findings operate through distinct mechanisms and in different sectors than the attribute-based gains identified in this paper. The inflation inequality literature relies on scanner data from consumer packaged goods—a domain where innovation takes the form of product proliferation rather than attribute improvement. By contrast, the quality, dematerialization, and risk-reduction channels of Ω_t operate primarily in durables, transportation, and digital services, where technological progress is embedded in components and regulatory standards that apply across price tiers. Moreover, cheapflation is a short-run phenomenon tied to asymmetric pass-through of cost shocks, not a structural trend. For the long-run comparison that motivates this paper, the two phenomena coexist without contradiction.

6.4 Distinguishing Power from Welfare

6.4.1 A Formal Framework for Welfare vs. Power Inequality

Let aggregate welfare be given by:

$$W_t = U(C_t, S_t), \tag{23}$$

where C_t denotes the flow of priced consumption services and S_t captures the consumer surplus arising from unpriced digital goods, risk reduction, and quality improvements.

To avoid confusion with the notation P_t used throughout the paper for nominal prices, we denote political and institutional power by Π_t. Formally, let

$$\Pi_t = V(X_t),$$

where X_t denotes the vector of institutional, regulatory, and media resources that determine discretionary authority.

Power is introduced here as an analytical distinction, not as a measurable object. The argument does not depend on constructing an index of X_t; it only requires noting that the capacity to influence collective decisions is conceptually distinct from material living standards and is not captured by deflated income measures.

Assume weak separability between welfare and power:

$$U(C_t, S_t) \;\perp\; V(X_t). \tag{24}$$

Technological change affects W_t primarily through the expansion of S_t:

$$\frac{\partial W_t}{\partial t} = U_C \frac{\partial C_t}{\partial t} + U_S \frac{\partial S_t}{\partial t}, \tag{25}$$

whereas shifts in wealth concentration affect Π_t through:

$$\frac{\partial \Pi_t}{\partial t} = V_X \frac{\partial X_t}{\partial t}. \tag{26}$$

The empirical observation of this paper can be stated succinctly:

$$\frac{\partial S_t}{\partial t} > 0 \quad \text{for the median consumer,} \tag{27}$$

driven by technological improvements, while:

$$\frac{\partial X_t}{\partial t} > 0 \quad \text{for top deciles only,} \tag{28}$$

driven by institutional and financial dynamics.

Thus, inequality in W_t has narrowed due to the democratization of quality, while inequality in Π_t has widened due to the concentration of influence. Confusing the two leads to erroneous interpretations of economic progress.

This analysis does not deny that inequality exists or that it matters. Rather, it reframes what inequality *means* in a modern economy.

Wealth and income confer two analytically distinct benefits:

1. **Consumption services**: the direct utility from goods, services, safety, and experiences that income enables.

2. **Power**: the capacity to influence decisions, shape institutions, direct resources, and affect the lives of others.

The first has been substantially democratized by technology. The second has not. A billionaire's capacity to fund political campaigns, acquire media outlets, endow foundations, and shape public discourse vastly exceeds that of a median worker. This is a real and consequential form of inequality.

But conflating these two dimensions—as the inequality literature routinely does—produces a distorted picture of social reality. The claim that "the gains from growth have accrued to the top 1%" is true for power and positional goods. It is substantially false for the quality of daily consumption.

The separability between welfare and power is a static property of current consumption: in the present period, households across the income distribution consume goods and services of comparable quality. Over longer horizons, however, power can influence the institutional environment and thereby shape future welfare. Separating the two dimensions clarifies the argument: the measurement bias concerns the present distribution of material living standards, while the dynamics of political influence affect the evolution of those living standards over time. Acknowledging this distinction makes the welfare analysis sharper without denying that power can purchase future advantage.

The political implications are significant. A narrative focused on *material deprivation*—the claim that ordinary people are getting poorer, that living standards are declining, that the middle class is being immiserated—is empirically weak. A narrative focused on *power* *concentration*—the claim that the capacity to shape collective decisions is increasingly unequal—is empirically strong. Conflating the two weakens the case for the second by yoking it to the implausible first.

6.5 Policy Implications

The mismeasurement documented in this paper has concrete policy implications:

6.5.1 Monetary Policy

If inflation is systematically overstated, then real interest rates have been higher than policymakers perceived. A central bank targeting 2% inflation as measured by the CPI may in fact be targeting 1% true inflation or less. This implies a tighter monetary stance than intended, with consequences for employment and output.

6.5.2 Fiscal Policy

Many government programs are indexed to the CPI: Social Security benefits, tax brackets, poverty thresholds. If the CPI overstates inflation, then:

- Social Security benefits have grown faster in real terms than intended, contributing to long-term fiscal pressure.
- Tax brackets have widened faster than intended, reducing effective tax rates.
- Poverty thresholds have risen faster than true living costs, potentially overstating measured poverty rates.

These are not arguments for cutting benefits or raising taxes. They are arguments for accurate measurement. Policy choices should be made on the basis of reality, not statistical artifacts.

6.5.3 The Inequality Debate

The most important policy implication concerns the framing of the inequality debate itself. If the narrative of material immiseration is false, then policies predicated on that narrative may be misdirected.

This does not mean inequality is unimportant. It means that the *form* of inequality that matters most is not the one that dominates public discourse. Policies aimed at redistributing measured income may do little to address the concentration of political and institutional power. Conversely, policies aimed at diffusing power—antitrust enforcement, campaign finance reform, institutional decentralization—may do more to address consequential inequality than tax-and-transfer schemes.

6.6 A Revised Narrative

The evidence presented in this paper suggests a revised narrative of economic progress in advanced economies:

1. Real living standards have improved substantially over the past half-century—far more than official statistics indicate.
2. This improvement has been broadly shared in terms of *quality of consumption*: access to information, communication, safety, and entertainment has converged across income levels.
3. The improvement has *not* been broadly shared in terms of *power*: the capacity to influence collective decisions has become more concentrated.
4. The dominant narrative conflates these two dimensions, producing a picture of material immiseration that does not match lived experience, while obscuring the real and consequential inequality of power.

This revised narrative is neither complacent nor utopian. It acknowledges genuine progress while identifying the forms of inequality that remain unaddressed. It rejects the nostalgia implicit in claims that ordinary people were better off in some imagined past. And it redirects attention from statistical artifacts to the substantive question: who decides?

7 Conclusion

This paper began with a paradox: advanced economies report stagnating productivity and sluggish measured wage growth, yet the lived experience of technological transformation suggests a different reality. We have argued that this paradox is largely an artifact of statistical obsolescence.

The System of National Accounts was designed for an economy of standardized physical goods rolling off assembly lines. It is structurally incapable of capturing many forms of value in an economy where goods undergo continuous qualitative transformation, where new categories of consumption emerge from nothing, and where services once reserved for the wealthy become universally accessible at zero marginal cost.

Summary of Findings

We developed an attribute-based framework that reconceptualizes economic output as a flow of services from evolving bundles of characteristics. This framework identifies three sources of systematic mismeasurement: (i) quality improvements, (ii) dematerialization, and (iii) risk reduction. Aggregating across sectors, we estimate that true living standards have grown approximately 0.8–1.5 percentage points per year faster than official statistics suggest. Cumulated over decades, this implies that real wealth is substantially higher than measured—perhaps 50 percent higher over a forty-year horizon.

Implications for the Inequality Debate

The findings challenge the dominant narrative of stagnation and immiseration. The claim that "real wages have stagnated since the 1970s" relies on deflating nominal wages by a price index that systematically overstates inflation and ignores the emergence of entirely new categories of consumption. The claim that "the gains from growth have accrued to the top 1 percent" conflates two distinct phenomena: the concentration of power and the distribution of welfare.

We do not deny that inequality exists. We argue that the inequality which matters most— the concentration of political and institutional power—is obscured rather than illuminated by the focus on measured income distributions. The billionaire and the median worker have access to the same Google, the same antibiotics, the same airbags. They do not have access to the same capacity to shape institutions or constrain collective decisions.

A narrative of material deprivation does not match reality and weakens the case for addressing the consequential inequalities of power that characterize modern economies. The revised narrative suggested by this paper acknowledges substantial and broadly shared progress in living standards while identifying the forms of inequality that remain consequential and unaddressed. A final refinement concerns distributional dynamics. While structural improvements in durables, safety, and digital services have been broadly shared, recent evidence of inflation inequality in consumer packaged goods indicates that some categories have moved in the opposite direction. These forces operate through different mechanisms, and a complete account of distributional welfare change will require integrating both.

Closing Reflection

We are not poorer than we think. We are measuring our wealth with instruments designed for a vanished economy, and the resulting statistics have become the foundation for a narrative of decline that bears little resemblance to the lives people actually lead. The task is to see clearly: to distinguish genuine problems from statistical artifacts, and to direct attention to the inequalities that matter rather than the ones that make for convenient political narratives. The wealth of nations is greater than we count. The question is whether we will use that wealth wisely—and who will decide.

Acknowledgements

In preparing this manuscript, the author used AI assistants (Claude, ChatGPT, and Gemini) as research and editorial tools. These tools assisted with literature searches, provided feedback on expository clarity and structural coherence, and helped refine the presentation of arguments. The author was responsible for conceiving the central thesis, directing all research inquiries, selecting and critically evaluating sources and evidence, developing the theoretical framework, and exercising full editorial control over the final text. All substantive intellectual judgments—including the interpretation of findings and the paper's conclusions—are entirely the author's own.

References

Mark Bils and Peter J Klenow. Some evidence on the importance of sticky prices. *Journal of Political Economy*, 109(5):1135–1171, 2001.

Michael J Boskin, Ellen R Dulberger, Robert J Gordon, Zvi Griliches, and Dale Jorgenson. Toward a more accurate measure of the cost of living. Technical report, Final Report to the Senate Finance Committee, 1996.

Erik Brynjolfsson, Avinash Collis, W Erwin Diewert, Felix Eggers, and Kevin J Fox. Gdp-b: Accounting for the value of new and free goods in the digital economy. *NBER Working Paper*, (w25695), 2019.

David M. Byrne, John G. Fernald, and Marshall B. Reinsdorf. Does the united states have a productivity slowdown or a measurement problem? *Brookings Papers on Economic Activity*, 2016(1):109–182, 2016.

Alberto Cavallo and Oleksiy Kryvtsov. Price discounts and cheapflation during the post-pandemic inflation surge. *Journal of Monetary Economics*, 148:103644, 2024. doi: 10.1016/j.jmoneco.2024.103644.

David M Cutler and Mark McClellan. Is technological change in medicine worth it? In *Medical Care Output and Productivity*. University of Chicago Press, 2001.

F. Erhun, R. S. Kaplan, V. G. Narayanan, et al. Are cost advantages from a modern indian hospital transferable to the united states? *American Heart Journal*, 224:148–155, 2020.

Robert J Gordon. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton University Press, 2016.

V. Govindarajan and R. Ramamurti. Delivering world-class health care, affordably. *Harvard Business Review*, 2013.

B. Gupta, R. Huckman, and T. Khanna. Task shifting in surgery: Lessons from an indian heart hospital. *Healthcare*, 3(4):245–250, 2015.

Jerry A Hausman. Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods*, pages 209–237. University of Chicago Press, 1996.

Xavier Jaravel. The unequal gains from product innovations: Evidence from the U.S. retail sector. *The Quarterly Journal of Economics*, 134(2):715–783, 2019. doi: 10.1093/qje/qjy031.

Dale W. Jorgenson. Capital theory and investment behavior. *American Economic Review*, 53(2):247–259, 1963.

R. S. Kaplan and S. R. Anderson. Time-driven activity-based costing. *Harvard Business Review*, 2004.

R. S. Kaplan and S. R. Anderson. The innovation of time-driven activity-based costing. *Journal of Cost Management*, 2007.

G. Keel, C. Savage, M. Rafiq, and P. Mazzocato. Time-driven activity-based costing in health care: A systematic review. *Health Policy*, 121(7):755–763, 2017.

Kelvin J Lancaster. A new approach to consumer theory. *Journal of Political Economy*, 74(2):132–157, 1966.

National Highway Traffic Safety Administration. Traffic safety facts 2019: A compilation of motor vehicle crash data. Report DOT HS 813 141, National Center for Statistics and Analysis, 2021. Available at https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813141.

William D Nordhaus. Do real-output and real-wage measures capture reality? the history of lighting suggests not. In *The Economics of New Goods*, pages 29–66. University of Chicago Press, 1996.

Ariel Pakes. A reconsideration of hedonic price indexes. *NBER Working Papers*, (8715), 2003.

Thomas Piketty. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, MA, 2014. Translated by Arthur Goldhammer.

B. Richman and K. Schulman. What u.s. hospitals can still learn from india's private heart hospitals. *NEJM Catalyst*, 2017.

Sherwin Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974.

Emmanuel Saez and Gabriel Zucman. Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2): 519–578, 2016.

Emmanuel Saez and Gabriel Zucman. The rise of income and wealth inequality in america: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, 34(4):3–26, 2020.

Chad Syverson. Challenges to mismeasurement explanations for the us productivity slowdown. *Journal of Economic Perspectives*, 31(2):165–186, 2017.

P. Tseng, R.S. Kaplan, B.D. Richman, et al. Administrative costs associated with physician billing and insurance-related activities. *JAMA*, 319:691–697, 2018.

U.S. Department of Energy. Refrigerator and freezer energy rating database. Technical report, Office of Energy Efficiency and Renewable Energy, 2021. Available at https://www.energy.gov/scep/wap/articles/refrigerator-and-freezer-energy-rating-online-search-tool.

Gabriel Zucman. Global wealth inequality. *Annual Review of Economics*, 11:109–138, 2019.

Appendices

A Foundations of the Attribute-Based Model

This appendix provides the formal underpinnings of the attribute-based framework introduced in Section 2. The main text develops the economic intuition and sectoral applications; here we establish the mathematical structure required for the decomposition of the adjustment factor Ω_t and for the aggregation of measurement biases across goods and time.

A.1 Attribute Space and Valuation Functions

Let each consumption good x_t at time t be represented as a vector of attributes:

$$x_t = \big(q_t, d_t, s_t\big), \tag{29}$$

where:

- $q_t \in \mathbb{R}_+^m$ denotes quality-enhancing characteristics (durability, efficiency, safety, reliability);
- $d_t \in \mathbb{R}_+^k$ denotes dematerialized functionalities (digital capabilities replacing physical goods);
- $s_t \in \mathbb{R}_+^r$ denotes risk-reducing characteristics (probability reductions in hazards associated with consumption).

Consumers derive utility from the *services* generated by attributes, not from the attributes themselves. We therefore introduce valuation functions:

$$v_q : \mathbb{R}_+^m \to \mathbb{R}_+, \tag{30}$$
$$v_d : \mathbb{R}_+^k \to \mathbb{R}_+, \tag{31}$$
$$v_s : \mathbb{R}_+^r \to \mathbb{R}_+. \tag{32}$$

We impose the following minimal assumptions:

1. **Monotonicity:** v_q, v_d, and v_s are strictly increasing in each argument.
2. **Weak separability:** The contribution of each attribute block to the service flow does not depend on the levels of the other blocks. This does not require additivity, only that a strictly monotone aggregator can be constructed.
3. **Normalization:** $v_q(q_0) = v_d(d_0) = v_s(s_0) = 1$ for a chosen base period $t = 0$.

These assumptions are sufficient to define growth factors:

$$\theta_t = \frac{v_q(q_t)}{v_q(q_0)}, \tag{33}$$
$$1 + \delta_t = \frac{v_d(d_t)}{v_d(d_0)}, \tag{34}$$
$$1 + \rho_t = \frac{v_s(s_t)}{v_s(s_0)}. \tag{35}$$

The valuation functions are not intended to generate behavioral predictions. Their role is purely decompositional, consistent with index-number theory (Laspeyres, Paasche, Divisia), where monotonicity and weak separability are sufficient for an internally coherent adjustment factor.

A.2 Intertemporal Comparability

Standard price indices require *intertemporal comparability*: a good x_t must be comparable to $x_{t'}$ for the deflator to be meaningful. Formally, this requires the existence of:

$$\phi_{t,t'} : \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^r \to \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^r \tag{36}$$

such that:

$$x_t \sim x_{t'} \quad \iff \quad \phi_{t,t'}(x_t) = x_{t'}. \tag{37}$$

For homogeneous goods with fixed attributes, $\phi_{t,t'}$ is the identity. For evolving goods, comparability requires that attribute changes can be mapped onto a stable valuation structure.

Lemma (Failure of Fixed-Basket Comparability) If $x_t = (q_t, d_t, s_t)$ evolves such that at least one component is strictly monotonic in t and unbounded, then no fixed basket index can preserve intertemporal comparability.

Sketch. Suppose q_t increases strictly with t. A fixed basket implicitly assumes $q_t = q_0$ for all t, contradicting monotonicity. If new digital attributes appear ($d_t \notin \text{span}(d_0)$), no mapping $\phi_{t,0}$ exists. If hazards collapse ($s_t \to s_\infty$), valuation becomes incomparable across periods. □

This establishes that the CPI and GDP deflator cannot measure real consumption correctly in the presence of sustained quality change, dematerialization, or risk reduction.

The sketch is sufficient for establishing non-comparability; more formal proofs follow directly from standard results in index-number theory and are omitted for brevity.

A.3 The Multiplicative Structure of Ω_t

Given weak separability, the overall service flow from x_t is represented by a strictly monotone aggregator:

$$V(x_t) = F\big(v_q(q_t), v_d(d_t), v_s(s_t)\big). \tag{38}$$

Under standard homotheticity assumptions:

$$F(a, b, c) = a \cdot b \cdot c, \tag{39}$$

yielding:

$$\Omega_t = \theta_t \cdot (1 + \delta_t) \cdot (1 + \rho_t). \tag{40}$$

This multiplicative form is natural because each attribute block contributes independently to the service flow. Additive alternatives:

$$\tilde{\Omega}_t = \alpha\theta_t + \beta(1 + \delta_t) + \gamma(1 + \rho_t) \tag{41}$$

are theoretically inferior because:

1. They violate dimensional consistency unless θ_t, δ_t, and ρ_t share common units.
2. They permit negative contributions under rebasing.
3. They generate counterfactual predictions for goods where one component dominates (e.g., digital goods with $\delta_t \gg 1$).

This multiplicative formulation of the adjustment factor is thus the minimal structure consistent with attribute-based valuation.

A.4 Aggregation from Micro Attributes to Macro Bias

Let i index goods with expenditure share ω_i. The effective price is:

$$P^{\dagger}_{i,t} = \frac{P_{i,t}}{\Omega_{i,t}}. \tag{42}$$

Measured real expenditure is:

$$Y_t = \sum_i \omega_{i,t} \frac{P_{i,t}}{P_{i,0}} Q_{i,t}, \tag{43}$$

while true real expenditure is:

$$Y_t^* = \sum_i \omega_{i,t} \frac{P_{i,t}}{P_{i,0}\,\Omega_{i,t}} Q_{i,t}. \tag{44}$$

Let $B_t = Y_t - Y_t^*$. By log-differentiation and first-order approximation:

$$\frac{\dot{B}_t}{Y_t} \approx \sum_i \omega_i \left(\frac{\dot{\theta}_{i,t}}{\theta_{i,t}} + \frac{\dot{\delta}_{i,t}}{1+\delta_{i,t}} + \frac{\dot{\rho}_{i,t}}{1+\rho_{i,t}} \right), \tag{45}$$

which reproduces the aggregate bias expression in Section 5.

This completes the formal derivation of the attribute-based measurement framework.

A.5 Shadow Pricing and the Effective Budget Constraint

This section provides a formal derivation of the mechanism discussed in the main text: how digital services reduce shadow prices and relax the effective budget constraint. The relationship between consumer surplus and real income can be made explicit by considering a standard consumption model augmented with transaction and information costs. Let q_t denote the quantity of a representative consumption good and p_t its posted price. In the presence of search frictions, uncertainty, and information costs, the household faces an *effective* budget constraint of the form:

$$p_t q_t + \tau_t(q_t) \;\leq\; y_t, \tag{46}$$

where $\tau_t(q_t)$ captures non-price costs associated with acquiring and using the good: search time, uncertainty about quality, misallocation errors, distance costs, and other transaction costs. The function τ_t is assumed to be increasing in q_t and strictly positive in the absence of digital information.

Shadow Price. The household behaves as if it faces an implicit "shadow price"

$$\tilde{p}_t \;=\; p_t + \frac{\partial \tau_t(q_t)}{\partial q_t}, \tag{47}$$

which incorporates frictions that raise the effective marginal cost of consumption above the posted price. The relevant constraint is therefore

$$\tilde{p}_t q_t \;\leq\; y_t. \tag{48}$$

Even with constant income y_t, a reduction in $\tilde{p}_t$ expands the attainable consumption set.

Digital Services as Friction Reducers. Let χ_t denote the stock of digital services available at time t—search engines, navigation tools, online reviews, real-time price comparison, and other dematerialized functionalities. We model frictions as:

$$\tau_t(q_t) \;=\; \tau(q_t, \chi_t), \qquad \frac{\partial \tau}{\partial \chi_t} < 0, \tag{49}$$

reflecting the fact that digital services reduce the non-monetary costs associated with consumption. Substituting (49) into (47), we obtain the comparative static:

$$\frac{\partial \tilde{p}_t}{\partial \chi_t} \;=\; \frac{\partial^2 \tau(q_t, \chi_t)}{\partial q_t \partial \chi_t} \;<\; 0. \tag{50}$$

Digital goods therefore reduce the shadow price of consumption, even though they do not increase nominal income y_t.

Implication for Consumer Surplus. Let V_t be the indirect utility obtained under the effective constraint (48). A first-order approximation yields:

$$\frac{\partial V_t}{\partial \chi_t} \;=\; -\,\lambda_t\, q_t \frac{\partial \tilde{p}_t}{\partial \chi_t} \;>\; 0, \tag{51}$$

where λ_t is the marginal utility of income. Digital services therefore expand the feasible consumption set by lowering shadow prices—a channel distinct from monetary income but equivalent in welfare terms.

Interpretation. Consumer surplus from digital goods cannot be used to pay rent, but it lowers the shadow cost of achieving a given level of consumption. The household cannot convert a mapping service into cash, but it can use it to reduce travel costs, avoid costly mistakes, locate cheaper retailers, and minimize waste. These adjustments shrink τ_t and hence $\tilde{p}_t$, relaxing the effective budget constraint even holding y_t fixed.

Formally, digital surplus is not income, but it plays an income-equivalent role by increasing real consumption possibilities. This mechanism is invisible in national accounts, which treat p_t rather than $\tilde{p}_t$ as the relevant price. The mismeasurement arises because the System of National Accounts records posted prices but not the frictions whose collapse expands real consumption possibilities.

B The Institutional Friction Model in Healthcare Services: Evidence from India and the United States

B.1 Introduction

This appendix provides the empirical and conceptual foundations for assessing how institutional frictions shape observed healthcare costs, using the structure developed in the main text as its analytical background.

As discussed in the main text, healthcare provides one of the clearest empirical tests for distinguishing genuine technological constraints from institutional frictions in service production. While Baumol's interpretation attributes rising relative costs in services to intrinsically limited productivity growth, the analysis in the paper shows that this view cannot account for the large and persistent divergence between technological progress in medical treatment and observed expenditure trends. The organizational features of Narayana Health have been examined extensively in the healthcare management literature, particularly with respect to task specialization, workflow engineering, and high-volume surgical delivery (Gupta et al., 2015; Richman and Schulman, 2017).

The purpose of this appendix is to examine this claim empirically by applying the Institutional Friction Model to a domain where technological capabilities are both well measured and internationally comparable. In particular, we draw on the comparative costing study of Erhun et al. (2020), which contrasts two U.S. hospitals with the high–volume cardiac surgery center Narayana Health (NH) in Bangalore. Because all three institutions employ equivalent medical technologies and meet comparable quality standards, differences in observed costs can be systematically decomposed into technological, organizational, and institutional components.

Time-Driven Activity-Based Costing (TDABC), the methodology employed by Erhun et al. (2020), is especially suited to this task: it quantifies the resource use associated with each clinical activity and separates non-transferable cost differences (such as wages and local input prices) from organizational drivers and institutional frictions.

Originally developed by Kaplan and Anderson (2004, 2007), TDABC has become a widely adopted framework in healthcare cost analysis, precisely because it establishes a transparent linkage between clinical activities and the underlying resource expenditures they require. This appendix uses that structure to map the empirical findings into the attribute-based framework developed in Section 2 and Section 3

The remainder of the appendix proceeds as follows. Section B.1 introduces the Institutional Friction Model and clarifies its relationship to the attribute-based representation of service output. Subsection B.2 summarizes the main features of TDABC relevant for decomposing cost drivers. Subsection B.3 presents the core evidence from Erhun et al. (2020). Subsection B.4 analyzes the resulting decomposition in light of the model. Subsection B.5 concludes with implications for the interpretation of Baumol's Cost Disease and for the measurement of real income. This structure mirrors the logic of the attribute-based framework itself, moving from theoretical formulation to empirical identification and finally to interpretation within the broader economic context.

B.2 The Institutional Friction Model

We now formalize the analytical structure used to interpret these institutional and organizational differences. The attribute-based framework developed in the main text represents the effective service flow from a good or activity as an evolving bundle of characteristics whose value grows through quality improvements, organizational innovation, and reductions in risk. These components are summarized by the multiplicative factor $\Omega_t = \theta_t(1+\delta_t)(1+\rho_t)$. In the absence of external distortions, improvements in these attributes translate directly into lower effective costs or higher measured output.

The Institutional Friction Model introduces a further component that operates outside the technological and organizational domain. Let P_t^{tech} denote the technologically determined unit cost of producing a given health service at time t, and let P_t^{obs} denote the observed market cost. Institutional constraints create a wedge between the two:

$$P_t^{\text{obs}} = \frac{P_t^{\text{tech}}}{\Omega_t}(1+\Phi_t), \tag{52}$$

where Φ_t is the institutional friction parameter. Here, P_t^{tech} denotes the counterfactual cost that would obtain under frictionless institutional conditions, and is therefore distinct from the observed nominal price P_t used in the main text. The distinction mirrors the separation between technological efficiency and institutional distortions emphasized in Section 2 and Section 3. This parameter captures the extent to which regulatory structures, legal exposure, administrative requirements, and market design prevent technological and organizational improvements from being fully transmitted to consumers as lower prices or higher measured productivity.

For analytical clarity, we decompose the friction term into its principal components:

$$\Phi_t = \Phi_{\text{licensing}} + \Phi_{\text{liability}} + \Phi_{\text{admin}} + \Phi_{\text{oligopoly}}.$$

The additive form is adopted for expositional convenience; since the components of Φ_t are small in proportional terms, a first-order approximation is sufficient. Scope-of-practice regulations ($\Phi_{\text{licensing}}$) restrict the ability to reassign tasks to less costly personnel; malpractice exposure ($\Phi_{\text{liability}}$) induces defensive practices; billing and compliance requirements (Φ_{admin}) increase administrative overhead; and market concentration or entry barriers ($\Phi_{\text{oligopoly}}$) limit competitive pressure. Empirical evidence underscores the magnitude of these frictions in the U.S. healthcare system. In particular, administrative and billing requirements impose substantial overhead costs on providers, with detailed TDABC studies documenting that a significant share of total expenditure is absorbed by insurance-related activities rather than clinical care (Tseng et al., 2018). This empirical pattern is consistent with a large Φ_{admin} component, which acts as a structural wedge between technological capability and observed costs.

Together, these frictions determine how much of the technological and organizational progress embodied in Ω_t is reflected in observed costs.

The comparative costing study of Erhun et al. (2020) provides a setting in which the effects of Φ_t can be isolated. Because Narayana Health and the two U.S. hospitals examined in their study employ comparable medical technologies and achieve similar clinical quality, differences in observed costs can be attributed to organizational factors and institutional frictions rather than to differences in technology or outcomes. This makes their setting particularly well suited well suited for the institutional-cost decomposition carried out below.

B.3 Time-Driven Activity-Based Costing (TDABC)

Time-Driven Activity-Based Costing (TDABC) provides a measurement framework that allows technological, organizational, and institutional components of healthcare costs to be separated in a transparent and empirically grounded way. The method assigns costs on the basis of two observable elements: the minutes of personnel and space required for each activity in a clinical pathway, and the associated capacity cost rate for each resource. Because both elements are directly measurable, TDABC avoids the aggregation and allocation distortions typical of conventional hospital cost accounting. The suitability of TDABC for healthcare applications has been demonstrated in numerous studies, which show that process-level costing provides substantially greater accuracy than traditional accounting allocations, particularly in complex clinical pathways (Keel et al., 2017). These findings reinforce the methodological premise underlying Erhun et al. (2020), namely that a granular mapping of activities and resource usage is essential for distinguishing technological, organizational, and institutional sources of cost variation.

In the application developed by Erhun et al. (2020), the TDABC procedure begins by constructing detailed process maps for an isolated multivessel CABG surgery. These maps identify all relevant clinical activities—including intraoperative steps, postoperative monitoring, use of operating rooms, intensive care units, and ward beds—and record the time required from each category of personnel (senior and junior surgeons, anesthesiologists, nurses, technicians) as well as the occupancy of physical spaces. The resulting activity times are based on direct observation and interviews with clinical staff across the participating hospitals.

For each resource, the capacity cost rate is computed as the total annual cost of making that resource available divided by its practical capacity, expressed in clinical minutes per year. Multiplying activity times by the corresponding capacity cost rates yields the direct personnel and space cost of each stage of the CABG episode. Summing across activities produces a comprehensive estimate of the direct cost of care, disaggregated by resource type.

A central feature of the methodology in Erhun et al. (2020) is the sequential standardization of input prices and practical capacities across hospitals. In the first standardization, all hospitals are assigned the same wage and space cost parameters, allowing differences attributable solely to local labor markets or regional price levels to be removed. In the second standardization, all hospitals are assigned the same practical capacity assumptions for personnel and space, eliminating differences arising from institutional or contractual norms governing annual working hours or space utilization.

After these two adjustments, any remaining differences in total CABG cost reflect organizational design, workflow structure, and institutional frictions rather than technological inputs or regional price effects. TDABC therefore provides a direct empirical bridge between the cost decomposition in Erhun et al. (2020) and the Institutional Friction Model introduced in Subsection B.1, isolating the portion of cost differences attributable to δ_t (organizational innovation) and to the friction term Φ_t.

B.4 Case Study: Narayana Health and U.S. Hospitals

We now apply this TDABC framework to the comparative costing study of Erhun et al. (2020). The comparative costing study of Erhun et al. (2020) analyzes the direct personnel and space costs of an isolated, non-urgent, multivessel CABG procedure across three Joint Commission–accredited hospitals: Narayana Health (NH) in Bangalore, The Heart Hospital Baylor Plano (THHBP) near Dallas, and Intermountain Medical Center (IMC) in Salt Lake City. All three sites employ comparable cardiovascular surgical technologies and achieve similar quality benchmarks, making them suitable for a structured decomposition of cost differences. As noted by Erhun et al. (2020), all three hospitals meet Joint Commission standards, and their risk-adjusted CABG outcomes fall within comparable quality tiers.

NH is a high-volume center performing approximately 4,000 CABG surgeries per year, with ten cardiovascular operating rooms and more than forty cardiovascular surgeons on staff. IMC and THHBP, by contrast, perform roughly 500 and 200 CABG procedures per year, respectively. NH's organizational structure and high-volume surgical model have also been examined in the management and health services literature. Prior studies emphasize the role of structured task delegation, narrow task definition, and workflow specialization in enabling high throughput without compromising clinical quality (Gupta et al., 2015; Richman and Schulman, 2017). These features form an important contextual background for interpreting the cost differentials observed in Erhun et al. (2020). Although all three hospitals share modern clinical equipment and standardized CABG protocols, they differ substantially in organizational structure, especially in the allocation of personnel time and operating room utilization. These comparisons concern the cost and quality of care conditional on access, and do not address institutional processes governing access itself.

Indexed cost comparison. Table II of Erhun et al. (2020) reports the relative personnel and space costs of a CABG episode after three sequential adjustments. In the unadjusted comparison (Panel A), NH's costs are approximately one-tenth of IMC's and one-sixteenth of THHBP's in the unadjusted comparison. After standardizing input prices by assigning IMC's wage and space-cost structure to all hospitals (Panel B), NH's cost rises to 84% of IMC's and 55% of THHBP's. A further standardization to IMC's practical capacity assumptions (Panel C) yields a final indexed cost of 93% relative to IMC and 67% relative to THHBP. These adjustments remove differences attributable to local labor markets and to institutional or contractual norms governing annual working hours and space availability.

Personnel activity times. Table III provides detailed evidence on intraoperative and postoperative personnel minutes. In intraoperative care, NH relies heavily on junior surgeons and junior anesthesiologists, while senior anesthesiologists contribute substantially fewer minutes per case than at IMC or THHBP. Mid-level providers and technicians also play a larger role in NH's operating rooms. Postoperative care displays a similar pattern: NH uses fewer registered nurse minutes per patient and a correspondingly larger amount of technician time. These differences reflect a systematic reallocation of tasks toward less costly personnel while maintaining senior supervision, consistent with the hospital's high-volume workflow design.

Decomposition of cost drivers. Table IV in Erhun et al. (2020) decomposes the personnel-cost differences between NH and the U.S. hospitals into three components: differences in input prices, differences in annual work hours (practical capacity), and differences in organizational design methods. After standardizing for wages and work hours, the residual gap between NH and IMC reflects two opposing forces: a 33% advantage attributable to NH's lower-cost skill mix and a 25% disadvantage attributable to its use of a larger total quantity of personnel hours per case. The net effect is an 8% residual productivity advantage for NH, arising entirely from organizational design rather than technological or regional price differences.

Summary. The evidence from Erhun et al. (2020) shows that, after adjusting for non-transferable factors such as regional input prices and institutional norms governing practical capacity, the remaining cost gap between NH and its U.S. counterparts arises from differences in workflow structure, skill mix, and operating room utilization. These components provide the empirical basis for interpreting the NH–U.S. comparison within the Institutional Friction Model, as developed in the next subsection.

B.5 Interpretation of the Decomposition

The decomposition reported in Table IV of Erhun et al. (2020) separates the personnel-cost differences between Narayana Health and the U.S. hospitals into three components: differences in input prices, differences in practical capacity, and differences in care organizational design. This subsection maps these empirical components into the analytical structure developed in Subsection B.1.

The first component reflects differences in local wages and space costs. Since these differences arise from regional labor markets rather than from technological or organizational factors, they lie outside the attribute-based representation of service production. In the notation of Subsection B.1, they affect neither Ω_t nor Φ_t, and are therefore not informative about underlying productivity differences.

The second component captures differences in annual work hours and space availability (practical capacity). As in Erhun et al. (2020), these differences are treated as non-transferable across hospitals and are therefore removed through standardization. Although such capacity norms may in part reflect institutional features of each health system, their treatment in the study leaves them outside the scope of the present decomposition.

The third component, consisting of differences in organizational design, corresponds directly to organizational and institutional elements of the model. The skill-mix advantage of NH (component IIIa) is a clear instance of organizational design: tasks are reassigned from senior personnel to junior staff and technicians under structured supervision, reducing the average cost per unit of clinical activity. In the attribute-based framework, such innovations increase the organizational component of Ω_t.

By contrast, the greater total labor hours used per case at NH (component IIIb) represent a countervailing effect within the same organizational domain. This pattern reflects the substitution of many lower-cost junior personnel hours for fewer high-cost senior hours, a design choice that reduces total cost despite increasing the aggregate time input. The net effect of the two subcomponents, however, is an 8% productivity advantage in favor of NH after controlling for input prices and practical capacity. This residual component thus reflects differences in workflow structure and operating-room utilization that are attributable to organizational design rather than to technological or regional price factors.

Finally, to the extent that some of the organizational practices observed at NH—such as the scope of task delegation or the structure of clinical supervision—are constrained in the U.S. context by licensing regulations, liability exposure, or administrative requirements, the corresponding portion of the residual can be interpreted as part of the institutional friction term Φ_t. Similar considerations emerge in comparative analyses of Indian and U.S. healthcare delivery models, which highlight how licensing rules, liability regimes, and administrative constraints limit the feasibility of transferring high-volume, task-shifted surgical workflows to the U.S. context (Richman and Schulman, 2017). This evidence supports the interpretation that a non-negligible share of the residual cost differential in Erhun et al. (2020) reflects institutional rather than technological or organizational factors. TDABC therefore provides a direct empirical identification of the organizational and institutional components of the cost gap, bridging the decomposition in Erhun et al. (2020) with the theoretical structure developed in this appendix.

B.6 Synthesis and Implications

The preceding decomposition provides the basis for evaluating the broader implications of institutional frictions for healthcare costs and for the interpretation of Baumol's Cost Disease. Rather than revisiting the numerical results, this subsection highlights the consequences of the residual component for understanding why observed healthcare costs may diverge from underlying technological and organizational improvements.

In particular, the organizational practices that drive NH's cost advantage—such as the systematic reassignment of tasks to junior personnel under structured supervision—illustrate forms of organizational design that are difficult to implement in the U.S. context. Licensing rules, liability exposure, and administrative requirements represent substantial institutional barriers, as noted in comparative analyses (Richman and Schulman, 2017).

These considerations are consistent with the Institutional Friction Model introduced in Subsection B.1: technological and organizational improvements raise Ω_t, but externally imposed constraints determine the extent to which these gains translate into observed cost reductions for patients and payers.

Similar patterns have been documented in other high-volume clinical environments with minimal institutional frictions, such as the Aravind Eye Care System, where workflow specialization and structured task delegation yield substantial reductions in cost while maintaining clinical outcomes. This parallel reinforces the broader empirical conclusion that, when institutional constraints are limited, organizational design alone can generate large and sustained reductions in the effective cost of complex clinical services (Govindarajan and Ramamurti, 2013). These parallels reinforce the interpretation that observed cost differences primarily reflect organizational and institutional factors rather than technological limitations. Overall, the evidence from Erhun et al. (2020) is consistent with the broader theme of the paper: when organizational characteristics are optimized and institutional frictions are low, the effective cost of healthcare services follows a deflationary trajectory despite rising technological capability. Conversely, when institutional constraints are binding, observed prices may diverge from underlying technological progress, giving rise to the appearance of a cost disease in sectors where productivity is, in fact, increasing.

C Quantifying Quality, Dematerialization, and Risk Improvements: An Empirical Illustration

This appendix provides an empirical illustration of the three components of the adjustment factor

$$\Omega_t = \theta_t(1+\delta_t)(1+\rho_t),$$

using representative data from domains where each component is particularly salient: (i) quality improvements in durable goods, (ii) risk reduction in transportation, and (iii) dematerialization in digital convergence devices. The objective is not to construct a complete quality-adjusted deflator—a task that would require comprehensive data across all consumption categories—but to demonstrate that each component of Ω_t is *measurable* using publicly available data, and that the resulting corrections are *quantitatively large* relative to conventional price measurement.

C.1 Quality Improvements in Durables: Energy Efficiency of Refrigerators

Historical data from the U.S. Department of Energy indicate that a standard refrigerator-freezer consumed roughly $E_{1979} \approx 1800$ kWh/year in the late 1970s, compared with approximately $E_{2020} \approx 450$ kWh/year for modern units of comparable internal volume (U.S. Department of Energy, 2021). Since refrigeration is an energy-intensive service, a natural quality index based on operating efficiency is:

$$\theta_{\text{fridge},t} = \frac{E_{1979}}{E_t}. \tag{53}$$

This yields:

$$\theta_{\text{fridge},2020} = \frac{1800}{450} = 4.0. \tag{54}$$

The interpretation is that a 2020 refrigerator delivers four times the refrigeration service per unit of energy cost as its 1979 predecessor. Since nominal prices have remained roughly constant in inflation-adjusted terms while this fourfold efficiency gain occurred, the *effective price* of refrigeration services—the cost per unit of cold storage adjusted for operating expenses—has declined dramatically.

Table 2: Refrigerator Energy Consumption and Quality Index

Year	Energy Use (kWh/year)	Quality Index θ_t
1979	~ 1800	1.00
2020	~ 450	4.00

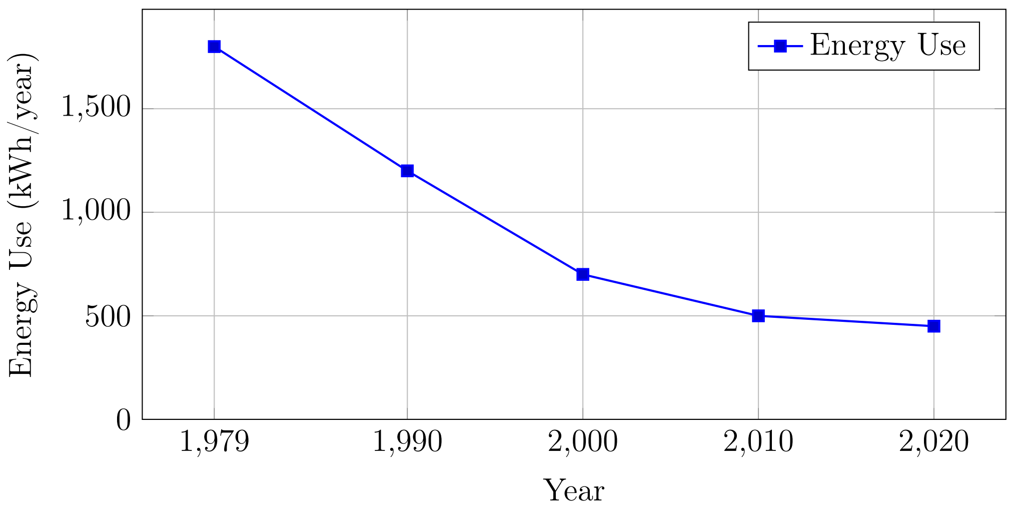

Figure 2: Decline in refrigerator energy consumption (1979–2020). Source: U.S. Department of Energy (2021).

This example isolates the θ_t component: the good has improved in a measurable quality dimension (energy efficiency) that directly affects the cost of obtaining the underlying service (refrigeration). Standard price indices, which track the purchase price of "a refrigerator," miss this improvement entirely.

C.2 Risk Reduction in Transportation: Fatality Rates

Automobile safety has improved dramatically due to engineering advances (crumple zones, airbags, electronic stability control) and regulatory requirements (NHTSA standards, mandatory seatbelt use). Fatality rates per 100 million vehicle miles traveled (VMT) in the United States fell from approximately 3.35 in 1980 to 1.11 in 2019 (National Highway Traffic Safety Administration, 2021).

A risk-reduction index can be constructed as the inverse hazard ratio:

$$1 + \rho_{\text{auto},t} = \frac{\pi_{1980}}{\pi_t}, \tag{55}$$

where π_t denotes the fatality rate at time t. This yields:

$$1 + \rho_{\text{auto},2019} = \frac{3.35}{1.11} \approx 3.0. \tag{56}$$

Table 3: Motor Vehicle Fatality Rates and Risk Index

Year	Fatalities per 100M VMT	Risk Index $1 + \rho_t$
1980	3.35	1.00
2019	1.11	3.02

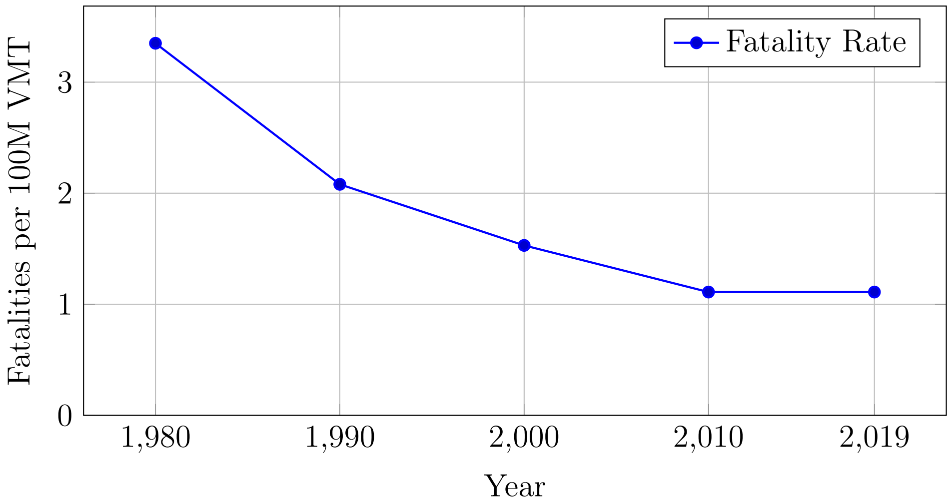

Figure 3: Motor vehicle fatality rate in the United States (1980–2019). Source: National Highway Traffic Safety Administration (2021).

The economic interpretation is that the *risk-adjusted* cost of transportation has fallen by a factor of three, even if the nominal cost of purchasing and operating a vehicle has remained stable or increased. Using standard value-of-statistical-life (VSL) estimates, this safety improvement represents an annual welfare gain on the order of hundreds of billions of dollars—a gain that appears nowhere in measured GDP or real income statistics.

C.3 Dematerialization: The Smartphone as Convergent Device

A modern smartphone replaces several categories of physical goods that were previously purchased separately. A conservative estimate of displaced expenditures, based on retail prices circa 2005, is shown in Table 4.

Table 4: Physical Goods Replaced by Smartphone Functionality

Functionality	Circa-2005 Cost (USD)
Digital camera (8+ MP)	250
GPS navigator	200
MP3 player (with storage)	150
Camcorder	350
Encyclopedia (CD/DVD set)	100
Voice recorder	50
Portable game console	150
Total displaced cost	**1,250**

Given an average smartphone price of approximately $800 in 2020 (for a device matching or exceeding the functionality listed), the dematerialization component is:

$$1 + \delta_{\text{phone},2020} = \frac{1250}{800} \approx 1.56. \tag{57}$$

This calculation captures only the *direct replacement* effect. It excludes the substantial consumer surplus from functionalities that were previously unavailable at any price: real-time global communication, instant access to the world's information, streaming media libraries, and sophisticated applications. Including these would raise the dematerialization index substantially, as suggested by the willingness-to-pay estimates in Brynjolfsson et al. (2019).

C.4 Interpretation and Limitations

The three examples illustrate distinct channels through which conventional price measurement understates welfare growth:

1. **Quality channel** (θ_t): Refrigerators deliver more service per dollar due to efficiency improvements invisible to price indices.

2. **Risk channel** (ρ_t): Automobiles are dramatically safer, a welfare gain with high monetary value that is unrecorded in national accounts.

3. **Dematerialization channel** (δ_t): Digital devices replace baskets of physical goods, generating both direct savings and consumer surplus from new functionalities.

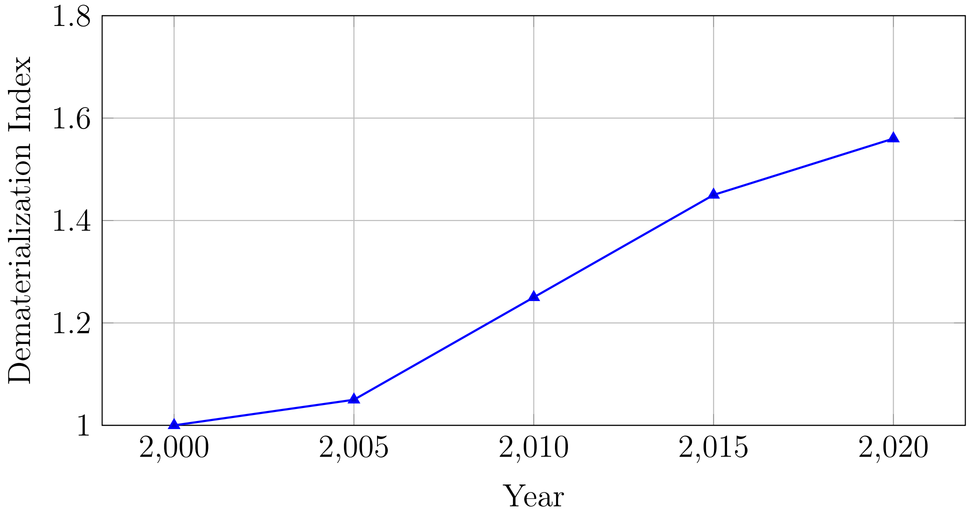

Figure 4: Growth of dematerialization through smartphone functionality.

Several limitations of this illustrative exercise should be noted. First, the examples are deliberately chosen for data availability and salience; they are not representative of the consumption basket as a whole. Second, we do not aggregate across categories, as doing so would require comprehensive attribute data and careful treatment of substitution patterns. Third, the calculations provide point estimates without formal uncertainty quantification.

Despite these limitations, the exercise establishes a central empirical point: the components of Ω_t identified in the theoretical framework are not abstract constructs but measurable magnitudes. In each case examined, the correction implied by proper quality, risk, or dematerialization adjustment is *large*—on the order of a factor of 1.5 to 4.0 for a single good. If similar patterns hold across the consumption basket—as the sectoral evidence in Section 4 suggests—the cumulative bias in measured real income growth is substantial.

The aggregate estimates reported in Section 5, drawing on the Boskin Commission and subsequent literature, are consistent with the magnitudes illustrated here. A complete implementation of the attribute-based framework, applied systematically across consumption categories, would likely yield corrections at least as large as those suggested by the illustrative calculations in this appendix.

C.5 Consistency with Aggregate Estimates

The point estimates derived above can be translated into implied annual growth rates to verify consistency with the aggregate bias estimates reported in Section 5.

For refrigerators, $\theta_{2020}/\theta_{1979} = 4.0$ over 41 years implies an annual quality growth rate of approximately $\ln(4.0)/41 \approx 3.4\%$ for this specific product. Weighted by the expenditure share of refrigeration services (approximately 0.5%), the contribution to aggregate bias is on the order of 0.02 percentage points per year. The threefold reduction in automobile

Table 5: Implied Annual Growth Rates by Component

Component	Level Change	Period	Implied Annual Rate
Quality (θ): Refrigerators	4.0×	1979–2020	$\sim 3.4\%$
Risk ($1+\rho$): Automobiles	3.0×	1980–2019	$\sim 2.8\%$
Dematerialization ($1+\delta$): Smartphones	1.56×	2007–2020	$\sim 3.4\%$

fatality risk implies a comparable annual rate of $\ln(3.0)/39 \approx 2.8\%$. For smartphones, the dematerialization index of 1.56 achieved over roughly 13 years (from the iPhone launch in 2007) implies $\ln(1.56)/13 \approx 3.4\%$ annual growth.

These product-level rates are high, but each product represents only a small share of total consumption. The key insight is that *similar patterns hold across broad categories of goods*—durables, transportation, digital services—each contributing incrementally to aggregate bias. When summed across the consumption basket, the cumulative effect yields the 0.8–1.5 percentage point annual bias reported in Section 5, consistent with the Boskin Commission and subsequent literature.

D Empirical Strategy, Data Sources, and Measurement Considerations

This appendix outlines the empirical foundations necessary to operationalize the attribute-based measurement framework developed in the main text and formalized in Appendix A. The goal is not to provide exhaustive estimates but to establish the methodological path by which the components of the adjustment factor $\Omega_t = \theta_t(1+\delta_t)(1+\rho_t)$ can be estimated using existing or plausibly obtainable data. This serves both as a blueprint for empirical replication and as an agenda for future research.

D.1 Empirical Identification Strategy

The decomposition of the effective price $P_t^{\dagger}$ requires identifying empirical proxies for the evolution of quality (θ_t), dematerialization (δ_t), and risk reduction (ρ_t). Each of these components corresponds to distinct observable attributes of goods and services.

Quality improvement (θ_t). Quality changes can be identified through:

- hedonic regressions linking attributes to prices;
- total cost of ownership (TCO) calculations;
- measurable performance improvements (energy consumption, lifespan, failure rates);
- safety upgrades (crash-test scores, engineering standards);

- regulatory changes capturing minimum quality thresholds.

Dematerialization (δ_t). The replacement of physical goods by digital functionalities is identified by:

- historical price series for displaced goods (cameras, GPS units, audio players);
- physical product catalogs (Sears, RadioShack, JC Penney);
- consumer willingness-to-pay and willingness-to-accept experiments for digital services;
- app-based service logs (frequency and intensity of usage);
- market disappearance of legacy industries.

Risk reduction (ρ_t). Risk-adjustment relies on:

- historical hazard rates (foodborne illness, transportation safety);
- value-of-statistical-life (VSL) calibrations;
- FDA, USDA, and CDC safety data;
- engineering reliability metrics;
- accident, recall, and failure-rate time series.

Each component contributes multiplicatively to Ω_t, so accurate measurement does not require perfect precision for any single factor.

D.2 Data Sources and Practical Measurement Constraints

The following datasets provide the empirical basis for estimating the adjustment components:

Consumer and Expenditure Data.

- Bureau of Labor Statistics (BLS) CPI microdata.
- Consumer Expenditure Survey (CEX).
- Nielsen scanner data (retail purchases).
- Public scanner datasets: Dominick's (Chicago) and IRI Academic Releases.

Quality and Durability.

- EPA energy-efficiency and appliance consumption data.
- IEA efficiency time series.
- NHTSA crash-test and fatality-rate data (vehicles).
- UL safety standard databases.
- Home construction code archives (insulation R-values, structural standards).

Dematerialization.

- Sears and RadioShack historical catalogs (physical goods).
- eBay historical price distributions for legacy devices.
- App usage logs and market penetration statistics.
- Internet Archive (Wayback Machine) for digital service versions.

Risk Reduction.

- CDC FoodNet and PulseNet datasets.
- USDA contamination and sampling archives.
- FAA aviation safety databases.
- OSHA workplace hazard records.

Limitations.

- attribute-level data are uneven across time and product categories;
- scanner data omit digital goods entirely;
- TCO data require reconstruction for older goods;
- quality and risk attributes are often latent and require proxies.

Despite these limitations, the broad empirical regularities documented in Section 4 are robust to data imperfections.

D.3 Operationalizing the Attribute Model

To compute θ_t, δ_t, and ρ_t, attributes must be translated into valuation functions v_q, v_d, and v_s as defined in Appendix A.

Estimating θ_t. Quality multipliers can be constructed using:

(i) Energy-adjusted service hours per dollar;

(ii) Lifespan improvements (MTBF, survival curves);

(iii) Performance indices derived from third-party testing;

(iv) Reductions in maintenance frequency or cost.

Estimating δ_t. Dematerialization requires:

- computing the historical cost of displaced goods $\sum_k P_{g_k,t_0}$;
- subtracting the cost of the convergent device $P_{n,t}$;
- estimating consumer surplus by experimental or revealed-preference methods.

Estimating ρ_t. Risk multipliers rely on:

$$1 + \rho_t = \frac{v_s(s_t)}{v_s(s_0)} = \frac{\text{expected hazard cost}(t_0)}{\text{expected hazard cost}(t)}. \tag{58}$$

Hazard costs are monetized using:

- direct medical costs;
- lost wages;
- VSL-based valuations for mortality risk.

D.4 Benchmarking and Sensitivity Analysis

Given the uncertainty in attribute measurement, sensitivity analysis is essential. Key considerations include:

- elasticity of demand in computing consumer surplus;
- alternative choke-price assumptions;
- alternative depreciation schedules (for TCO adjustments);
- alternative hazard monetization values (VSL ranges);
- substitution patterns affecting expenditure weights ω_i.

A robust feature of the framework is that the signs of θ_t, δ_t, and ρ_t are unambiguously positive in sectors where innovation is substantial; only magnitudes vary.

D.5 Estimating the Aggregate Bias B_t

Let i index goods and services in the consumption basket. The aggregate bias is:

$$B_t = \sum_i \omega_{i,t} \left(P_{i,t} - \frac{P_{i,t}}{\Omega_{i,t}} \right), \tag{59}$$

with growth rate approximated as in Section 4:

$$\frac{\dot{B}_t}{Y_t} \approx \sum_i \omega_i \left(\frac{\dot{\theta}_{i,t}}{\theta_{i,t}} + \frac{\dot{\delta}_{i,t}}{1+\delta_{i,t}} + \frac{\dot{\rho}_{i,t}}{1+\rho_{i,t}} \right). \tag{60}$$

Estimates in the main text (0.8–1.5 pp/yr) can be reproduced by using:

- appliance and energy-efficiency data for θ_t,
- digital displacement lists for δ_t,
- food and transportation hazard declines for ρ_t.