



21 April 2026

Matthew Clark  
Manager, Price-Quality Regulation  
Commerce Commission  
P O Box 2351  
Wellington  
By email to [regulation.branch@comcom.govt.nz](mailto:regulation.branch@comcom.govt.nz)

Dear Matthew

**Re: Revised Draft GDB Default Price-Quality Path Determination.**

This following submission is in response to the Commerce *Revised Draft GDB Default Price-Quality Path Determination - Changes related to proposed amendments to input methodologies for gas distribution services – Draft decision reasons paper*, (draft reasons paper) and attachments, dated 2 April 2026, but more generally addresses the deeper concerns in the Commission’s DPP4 Draft Decision.

While I have represented MGUG in this process, and I am fully briefed on their arguments, this submission is separate from (but references) MGUG submissions. The submission addresses fundamental technical and institutional issues with the Input Methodologies (IM), and Weighted Average Cost of Capital (WACC) determinations, and the Commission’s lack of accountability for outcomes that matter to consumers. This is not intended as a criticism of people in the Commission, rather it addresses the design issues with the regulatory framework itself that lead it to outcomes intuitively at odds with real world experiences.

The Commission’s approach to PQ settings for gas pipelines is largely formulaic, based on input data (including IM and WACC determinations) with supporting models (, building blocks, financial) that drive the outputs (starting price). This is not new. The Building Block/ RAB/ WACC regulatory framework was developed in 2010, and has been used for setting prices in every regulatory period since. Moreover, this framework, which is overwhelmingly a product of the neoliberal reform era of the 1980s–1990s, can claim broad support as it has been extensively adopted by economic regulators in the wider anglosphere. New Zealand has arguably the most explicitly codified version of this regime. The historical precedent and broader coalition of regulators using this is however not evidence that the framework itself is sound. On the contrary, a deeper dive into the origins and foundations of orthodox micro-economics and institutional design reveal that ideology, not science, has shaped the framework which is assumed as settled, but needs an urgent and critical re-examination.

The first issue discussed is the regulatory framework itself. Its foundation starts with a false model of the economy, adopts empirically disproven theory, and results in real world outcomes inconsistent with the intent of Part 4 of the Commerce Act. The Commission would be remiss in its obligations to

protect consumer outcomes if it did not seriously reflect on the possible technical weakness of its framework to deliver the consumer welfare outcomes that it promises.

The second challenge is fundamentally institutional. The Commission is currently not strongly accountable for outcomes. Rather it appears to be judged on process quality, and evaluation is more strongly based on whether the Commission followed correct procedures, consulted appropriately, applied stated methodologies consistently, and made legally defensible decisions. None of this requires the Commission to demonstrate that actual consumers were made better off. In light of real and wider welfare losses being experienced by households and businesses from Commission decisions affecting energy transport charges in particular, this needs an urgent shift away from process quality to outcome-based welfare testing. This requires a different conception of regulatory accountability — one that accepts that a technically correct price control process can still produce harmful welfare outcomes, and that the regulator has a responsibility to detect and respond to those outcomes rather than simply defending the integrity of the process that produced them.

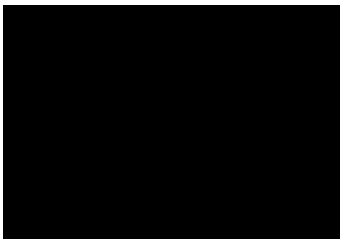
This is ultimately a governance question as much as a technical one, and it is worth noting that it sits in direct tension with one of the other legacies of the neoliberal reform era: the design of regulatory agencies as procedurally accountable but outcome-insulated institutions, shielded from political pressure precisely by being evaluated on process rather than result.

This submission is in two sections.

The first part deals with the theoretical foundations of the regulatory framework and exposes the fundamental flaws that has shaped it. This cannot be resolved for the DPP4 final decision, but awareness that it does not create outcomes that support consumer welfare can be used to moderate the high price rises and the distributional impacts being contemplated.

The second part deals with the topic of welfare outcomes and how to better control for them. Again, this is not an area that can be fully implemented in time for the DPP4 final decision, but the data and information already exist to allow the Commission to start applying better measures for assessing and moderating the impact of its DPP4 final decision.

Yours sincerely



Len Houwers  
Director Arete Consulting Ltd

## SUMMARY

- X 1 The underlying intent of Part 4 is social welfare maximisation (long-term benefit of consumers).
- X 2 The Building Block/ RAB/ WACC regulatory framework has its theoretical grounding in neoclassical economics, a dominant economic paradigm labelled as “orthodox”. Based on orthodox theory, social welfare is maximised when goods and services are delivered by competitive markets (s52A). Markets are deemed competitive when firms can only achieve “normal” profits (NPV=0).
- X 3 The foundations of neoclassical theory are however based on a false model of the economy. The economy is characterised by Complex Adaptive System (CAS) behaviour. Neoclassical economics however, implicitly assumes the economy is a closed, ergodic<sup>1</sup>, mechanical equilibrium system populated by homogeneous calculators interacting only through price signals in a world of decreasing returns and exogenous preferences, tending toward a unique, stable, Pareto-optimal equilibrium. Every one of these assumptions is a direct contradiction of CAS properties. The current framework does not fail to capture CAS dynamics by oversight — it is structurally constructed to exclude them. If theory starts with the wrong system analogy, then this should cast immediate doubt on the theory that follows from it.
- X 4 The neo-classical theory that follows from adopting the wrong system description, is equally disconnected from reality. The conclusions from neoclassical theory that provides for the regulatory policy levers are based on restrictive assumptions that do not reflect real world behaviour used to justify regulatory settings, including accelerated depreciation<sup>2</sup>:
- a. Equity losses are not borne by consumers, but by shareholders.
  - b. Risk transfer is not reversible. Intertemporal neutrality does not exist, and intertemporal shifting is not welfare neutral.
  - c. Distributional impacts cannot be averaged to a single representative agent, assumed to be the “average residential household”. Consumers do not have identical preferences, income levels, and behave the same way when averaged over the long term.
  - d. Consumers are not left “whole” when firms recalculate NPV=0 for each regulatory period.
  - e. Consumer behaviour is not rational, individual preferences are not stable, market demand curves can be any shape, marginal productivity curves are predominantly flat or decreasing, and market price is not an equilibrium outcome from intersecting supply and demand curves, but an administrative outcome.

---

<sup>1</sup> The assumption that a single representative agent experiences the same average outcome as many players.

<sup>2</sup> Most of these conclusions are explained in detail in MGUG’s submissions and cross submissions.

- f. Regulated asset owners are not single-purpose firms operating in a single market. When the actual firm is a diversified enterprise with regulated and unregulated subsidiaries, shared assets, consolidated financing, and group-level tax management, the model's outputs are an approximation at best and systematically biased at worst.
- g. The standard regulatory welfare test — does the price control deliver efficient prices that a hypothetical welfare-maximising consumer would prefer, is a fiction that is analytically convenient but empirically false. It means regulators can declare a price control "welfare-improving" while simultaneously causing serious harm to identifiable groups of actual consumers.

X 5 The Commission should adopt a real-world assessment of its modelled welfare outcomes for DPP4 rather than rely formulaically on its theoretical conclusions. The Commission should create an accountability mechanism that is entirely absent from current regulatory frameworks, most of which assess the quality of the process of price setting (consultation, methodology transparency, legal compliance) without ever measuring whether the outcome served consumer welfare as promised.

- a. Develop an analytical framework that identifies relevant dimensions of heterogeneity: income deciles, geographic location (urban vs rural, network periphery vs core), household composition (single adults, families with children, elderly households), tenure type (owner-occupiers vs renters, where the person paying the bill may differ from the person bearing the welfare impact), business size.
- b. For each segment, the Commission should model and then measure the share of household income or business cost absorbed by the regulated price, the responsiveness of consumption to price changes (low-income households typically have much lower price elasticity for essential services because they have already minimised consumption), and the availability of substitutes.
- c. A direct welfare measure for essential services is the expenditure share. Tracking this across the income distribution reveals the regressive or progressive incidence of regulatory decisions. Track this from DPP1 to test whether successive price controls have improved or worsened the distributional incidence of regulated prices.
- d. Establish explicit affordability thresholds —as binding outcome measures that the price control is required to meet or improve. If a threshold proportion of low-income households must spend no more than a defined share of income on the regulated service, this becomes a welfare constraint on the optimisation, not just a consideration.
- e. Measure actual quality outcomes by consumer segment rather than system-wide averages is essential to identify distributional harm.

- f. Rather than modelling what consumers *should* prefer based on assumed preferences, the Commission could measure what consumers *actually do* in response to regulatory changes<sup>3</sup>.
- g. For essential services with inelastic demand, price increases that exceed household budgets manifest not in reduced consumption but in non-payment, debt accumulation, and hardship programme uptake. These are highly direct welfare harm indicators that should be tracked systematically.
- h. Conduct spatial welfare analysis - Regulated networks are inherently spatial — the cost of serving a consumer depends critically on location, and uniform national or regional tariffs socialise these cost differences across the consumer base. This creates systematic cross-subsidies that have welfare implications for both those who pay above their cost-reflective share and those who are subsidised. Spatial welfare analysis would map the relationships (e.g., actual cost to serve by geographic unit, prices charged by geographic unit, consumer income and vulnerability by geographic unit).
- i. Conduct independent consumer welfare audits. An independent body would assess whether the price control as implemented has delivered the welfare outcomes the regulatory decision projected, identify which consumer segments have experienced welfare gains or losses, examine the distribution of actual returns to investors relative to what the WACC calculation allowed, and make findings publicly available to inform the next regulatory period.

---

<sup>3</sup> For example, if low-income households reduce consumption of an essential service following a price increase, this is direct evidence of welfare harm — households are being pushed below their preferred consumption level. For services with some degree of consumer choice or switching (where applicable), tracking switching behaviour, complaint rates, and hardship scheme uptake provides additional revealed preference signals. A spike in payment difficulties following a price control reset is direct evidence that the price path has exceeded the welfare tolerance of a portion of the consumer population

## CHALLENGING THE ECONOMIC PARADIGM

### Introduction

1. It might seem presumptive to engage academic economists at the Commission in a theoretical debate on an economic orthodoxy embedded in its structure and thinking, but the legitimacy to do so arises from within the academic discipline itself, and this submission draws on academic literature to make the case.
2. My expectation is that the Commission might reflect on this as possibly new information, since from personal experience, academic courses in economics generally do not teach students economic history, or critique foundations of the neoclassical theory taught exclusively at universities.
3. A number of theoretical challenge areas have already been covered in MGUG's submission and cross submission, which I authored for MGUG. Judging by the fact that the Commission is only consulting on an amendment to the IM to allow GDBs a new price reopener mechanism for an abnormal demand shock, rather than a full re-evaluation on its draft reasoning, my assumption is that the Commission has either dismissed or ignored these fundamental challenges to its framework. The reasons can only be speculated on since the Commission rarely engages in a debate on its reasoning. Paradigms however can be hard to shift and if these are being continually reinforced by the discipline in other jurisdictions, they become self-reinforcing. Equally possible is that the Commission is not explicitly accountable for outcomes of its decisions, but rather considers that it is primarily accountable for quality of the process of price setting (consultation, methodology transparency, and legal compliance), without ever measuring whether the outcome served consumer welfare as promised.
4. Nevertheless, I hope to be more convincing by addressing the actual foundation of Part 4 regulation to show where it is fundamentally flawed, and why the outcomes of the Commission's decisions in DPP4 are unlikely to be welfare promoting.
5. I finish by suggesting what the Commission could do to have a better view on whether its decisions actually improve or worsen social welfare outcomes.
6. This is all raised in the context of real-world concerns with escalating delivered energy costs leading to firm closures, output reductions, and increasing residential hardship. The Commission can't be oblivious to this deterioration in welfare, or the role it has in it. Commission representatives attended the manufacturers forum in 2025 organised by Optima Energy where they listened to the existential crisis facing small to medium manufacturers because of high and increasing delivered energy costs. MGUG has frequently explained how the distributional impacts of Commission pricing decisions in its submissions, particularly against the decision to accelerate depreciation, expose the weakness of the representative consumer model for assessing impacts. Nor could the Commission have ignored the more recent negative publicity on rising power prices attributed to mainly increases in line charges<sup>4</sup>. The case to reflect on and mitigate structural flaws is urgent.

---

<sup>4</sup> <https://www.rnz.co.nz/news/business/586471/consumer-nz-issues-power-price-warning>

## WRONG SYSTEM METAPHOR

7. Before commenting on Part 4 and IMs specifically, it is worth starting with the economic framework of neoclassical vs observed economic system behaviour to understand where the root of the problem appears to lie.
8. While not stated explicitly anywhere, the language used, and the methodological foundations of Part 4 PQ regulation show a strong adherence to neoclassical economic theory, especially micro-economics, which includes the branch of welfare economics. Neoclassical theory is the dominant economic paradigm embedded in the anglophone sphere including its institutions; universities, Treasuries, Reserve Banks, and economic regulators such as the Commission<sup>5</sup>.
9. The founders of neoclassical economics emerged in the Marginal Revolution of the 1870s, with three independent and near-simultaneous contributions from William Jevons<sup>6</sup>, Léon Walras<sup>7</sup>, and Carl Menger<sup>8</sup>. Later systematisers include; Alfred Marshall<sup>9</sup>, Francis Ysidro Edgeworth<sup>10</sup>, Vilfredo Pareto<sup>11</sup>, and Irving Fisher<sup>12</sup>
10. Some of their foundational theories and the empirical evidence against them are described in MGUG” cross submission<sup>13</sup>.
11. It is perhaps notable that the neoclassical foundations laid down in the 19<sup>th</sup> century was developed at a time where system dynamics was not understood, economic data was sparse, the science of applied statistics was only just emerging, and computing power was essentially manual. Theory was developed mostly from thought experiments heavy on simplifying assumptions to make them mathematically tractable.
12. At its heart, neoclassical theory uses Newtonian mechanics/ mechanical equilibrium system as its metaphor. This was not accidental - it was deliberate and is historically traceable. Mirowski (1989, *More Heat Than Light*, Cambridge) demonstrated that the neoclassical revolution of the 1870s directly imported the mathematical structure of mid-19th century classical mechanics — specifically the potential energy field — into economics. Utility became the economic analogue

---

<sup>5</sup> Neoclassical theory is not universally accepted or adopted globally. A notable exception is China which has stronger roots in classical economists such as Adam Smith and Karl Marx and which has synthesised its own theoretical basis, formally designated as a socialist market economy. Nordic countries also operate on social democratic principles with a stronger basis in Keynesian economic theory.

<sup>6</sup> Published *The Theory of Political Economy* (1871) and introduced marginal utility as the basis of value, replacing the classical labour theory of value, and applied mathematical and utilitarian reasoning to economic behaviour.

<sup>7</sup> Published *Éléments d'économie politique pure* (1874). Developed general equilibrium theory — the idea that all markets clear simultaneously, and laid the mathematical foundations that dominate modern microeconomics.

<sup>8</sup> Published *Grundsätze der Volkswirtschaftslehre* (Principles of Economics) (1871). Developed marginal utility independently; founded the Austrian School; emphasised subjectivism and individual action over mathematical formalism.

<sup>9</sup> *Principles of Economics* (1890) — synthesised and popularised neoclassical ideas; supply/demand curves, consumer surplus, partial equilibrium

<sup>10</sup> Indifference curves, mathematical economics

<sup>11</sup> Pareto optimality, ordinal utility

<sup>12</sup> Capital theory, interest rates, quantity theory of money. Interestingly Irving Fisher disowned most of his theories after being bankrupted in the Great Depression when his theory of finance which assumed permanent equilibrium in credit markets and no bankruptcies was demonstrably disproved.

<sup>13</sup> <https://www.comcom.govt.nz/assets/Documents/2026-gas-default-price-quality-path/MGUG-Cross-submission-on-Gas-DPP4-draft-decisions-12-February-2026.pdf> Appendix 1

of potential energy; maximising utility became the analogue of finding the minimum energy state; equilibrium prices became the analogue of mechanical equilibrium. Mirowski's argument is that Jevons, Walras, and Edgeworth did not derive their mathematical structure from observation of economies — they copied it from the physics of their era. The economy was assumed to have the mathematical structure of a potential field because potential fields were the respectable mathematical objects of the 1870s. The assumptions followed from the mathematics, not the other way around.

13. The implications of this foundational choice are enormous:
  - a. **Forces balance at equilibrium.** Supply and demand are forces that balance; the price is the equilibrium point. This is structurally identical to a mass on a spring.
  - b. **The system has a natural resting state.** Disturbances cause deviations, but the system returns to equilibrium — it is self-correcting by construction.
  - c. **The equilibrium is unique and stable** (in the basic versions). General equilibrium existence proofs (Arrow and Debreu 1954) established existence, but uniqueness and stability are much harder — the Sonnenschein-Mantel-Debreu theorem (1972–74) showed that essentially any aggregate excess demand function is consistent with general equilibrium, meaning the system need not converge to a unique stable equilibrium. ***Neoclassical practice largely ignored this result.***
14. In contrast the foundational framework of modern heterodox theory is the biological/ecological metaphor. The economy is seen as a nested complex adaptive system (CAS)<sup>14</sup>. The characteristics of a CAS are mapped against the evidence for the economy being best described as CAS (*Figure 1*).
15. There is a further CAS nuance, specific to the economy, which is institutional embedding. Unlike biological CAS (ecosystems, immune systems), the economy is embedded in institutions — legal systems, property rights regimes, central bank frameworks, and economic regulators like the Commerce Commission — that are deliberately designed and impose order on what would otherwise be more chaotic dynamics. This makes the economy a ***partially designed CAS***, which is somewhat unusual in the CAS taxonomy. North (1990) and Ostrom (1990)<sup>15</sup> both emphasise that institutions are themselves endogenous emergent phenomena, but they also constrain and channel the CAS dynamics in ways that pure ecological CAS do not experience. In other words, institutions are a force that acts on (perturbs) the CAS and an agent within the CAS<sup>16</sup>.

---

<sup>14</sup> Financial markets, labour markets, product markets, and macroeconomic dynamics are all CAS, but they operate at different speeds, with different agent populations, and different feedback structures. They interact with each other, creating a CAS-of-CAS structure that is even harder to analyse than a single-level system.

<sup>15</sup> North (1990, *Institutions, Institutional Change*); Ostrom (1990, *Governing the Commons*)

<sup>16</sup> Not that this is recognised by the institutions themselves who see themselves as controlling rather than steering agents.

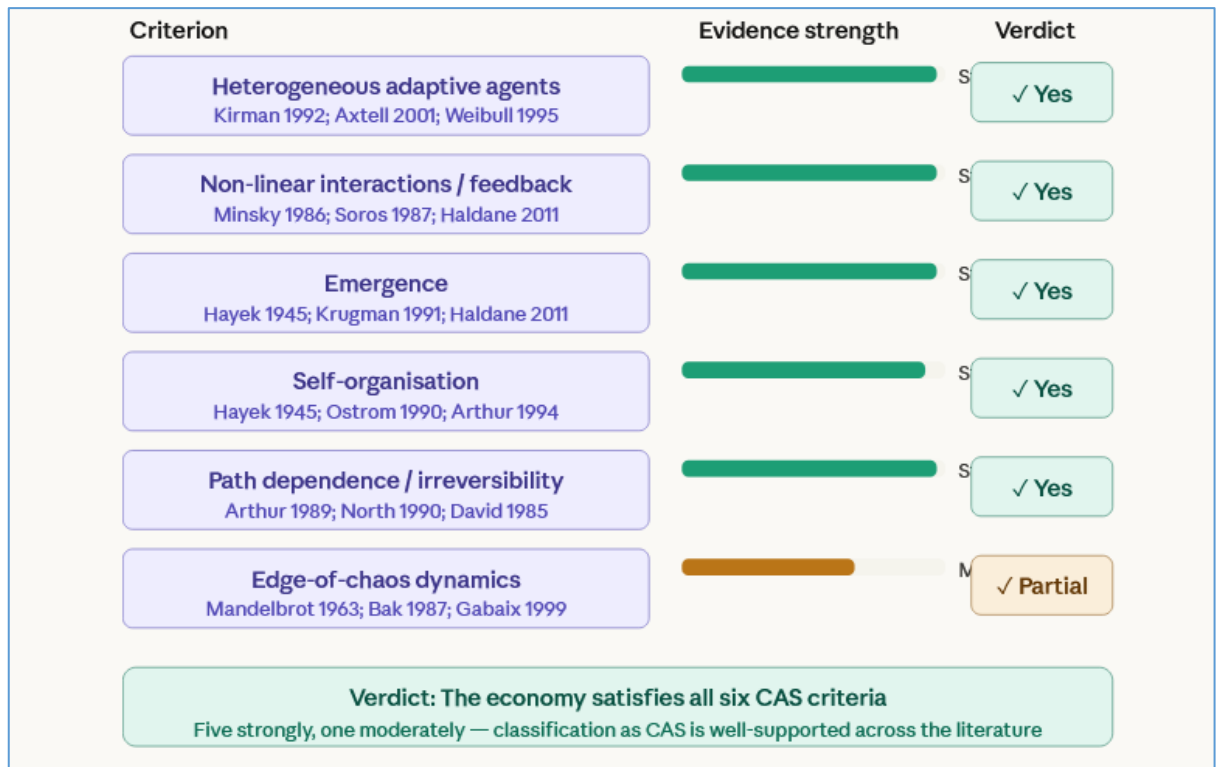


Figure 1: The economy as a CAS

16. The distinction between a mechanical equilibrium system and a CAS is critical to understanding what is likely to work in intervening vs what does not. The implications for how to analyse and intervene in the economy is shown in *Table 1*.

Table 1: Analysis and Intervention techniques

Mainstream framework	CAS-consistent framework
Find the equilibrium	Map the attractor landscape
Optimise policy parameters	Probe for leverage points (Meadows 1999)
Forecast specific outcomes	Scenario analysis with bounded distributions
Representative agent welfare	Distributional heterogeneity explicit
Treat crises as external shocks	Model crises as endogenous phase transitions
Comparative statics	System dynamics and path tracing

17. Table 1 demonstrates the Commission as firmly embedded in the mainstream framework. It also demonstrates MGUG’s criticism in their submissions of the Commission’s approach in both DPP3 and DPP4, particularly around forecasting, agent welfare, and crises treatment. The core insight from the table is that in a CAS, you cannot control the system; you can only perturb it and observe how it responds, influence the conditions under which agents adapt, and shift the attractors the system moves toward.

18. The deepest assumption that neoclassical theory proposes is that the economy is knowable in equilibrium terms. Behind all the specific assumptions is a meta-assumption that Keynes (1936) and Knight (1921) both identified: that the future is risky (probability-distributable) rather than uncertain (not reducible to probability distributions). Neoclassical agents maximise expected utility over a probability distribution of future states. This requires that the distribution is known or learnable — that the economy is the kind of system whose future states can be characterised by a stable probability distribution. In a CAS, this is false. Fundamental uncertainty — Knightian uncertainty<sup>17</sup> — is pervasive because the system's structure is continuously evolving, new agents and strategies emerge, and the distribution of future outcomes is not stationary. Shackle (1972, *Epistemics and Economics*) made this argument most forcefully: neoclassical economics is a theory of a closed system where all possibilities are already listed, probabilities are assignable, and the future is in principle calculable. A CAS is an open system where genuinely novel possibilities emerge that were not in any prior probability distribution. ***The neoclassical framework assumes the former; economies are the latter.***
19. This observation also lies at the heart of MGUG's objection to the Commission assuming its view on the outcome of the gas market into its decisions that has led to decisions such as accelerated depreciation.
20. In summary, neoclassical economics implicitly assumes the economy is a closed, ergodic, mechanical equilibrium system populated by homogeneous calculators interacting only through price signals in a world of decreasing returns and exogenous preferences, tending toward a unique, stable, Pareto-optimal equilibrium. **Every one of these assumptions is the precise negation of a CAS property. The framework does not fail to capture CAS dynamics by oversight — it is structurally constructed to exclude them.**
21. Ironically, CAS is also an explanation of why neoclassical theory is dominant through institutional path dependence. Once the framework was established, graduate education, journal publication standards, tenure criteria, and central bank hiring all selected for economists who could work within it. Heterodox approaches — Post Keynesian, evolutionary, institutional — produced results that were harder to formalise and thus less publishable in top journals. The selection environment of the profession reinforced the framework's own assumptions. This is itself a CAS dynamic: **the economics profession is a CAS that selected for the meme of neoclassical methodology**<sup>18</sup>
22. It should be clear from this *empirically tested and proven evidence* that economic regulation, such as applied by the Commission already starts from a fundamentally flawed system analogy. This is not a criticism particular to the Commission. The problem is endemic through

---

<sup>17</sup> Knightian uncertainty, coined by economist Frank Knight in 1921, distinguishes between measurable "risk" (known probabilities) and unquantifiable "uncertainty" (unknown probabilities or outcomes). It represents "unknown unknowns" where historical data cannot predict future events, making it impossible to calculate risks for decisions.

<sup>18</sup> The same selection pressure is what has created precise policy prescriptions within governments and central banks. Governments and central banks wanted clear, quantifiable policy advice — "raise rates by 25bp to reduce inflation by X%." CAS-based advice — "this is a non-linear system, we suggest probing with small interventions and observing" — is epistemically honest but institutionally inconvenient. The demand for precise policy prescriptions created selection pressure for frameworks that could produce them, regardless of their empirical validity.

almost all of the globalised world which has adopted neo-liberalism<sup>19</sup> and New Zealand was an enthusiastic adopter of this under Rogernomics. Nevertheless, the implications of this fundamental belief system reach deep into the design and implementation of Part 4.

### Where the IMs fail in its objective

23. Part 4 purpose statement is explicit;

*The purpose of this Part is to **promote the long-term benefit of consumers** in markets referred to in section 52 **by promoting outcomes that are consistent with outcomes produced in competitive markets** such that suppliers of regulated goods or services (etc)*

24. It is important to understand the context of the legislation. It was developed at a time of neo-liberal reform in New Zealand and leans exclusively on neoclassical economic theory to define the means to outcomes. The question here is why are competitive markets assumed to deliver long term benefits for consumers?

25. This core claim rests on two pillars formalised in welfare economics. The first fundamental theorem of welfare economics states that competitive markets, *under certain assumptions*, lead to *Pareto efficient outcomes*. The second theorem states that, *with further restrictions*, any Pareto efficient outcome can be achieved through a competitive market equilibrium, *given appropriate lump-sum transfers*.

26. A particularly interesting and implausible restriction is that the theorem requires **that a (central) social planner can redistribute endowments via lump-sum transfers** — taxes or grants that do not distort anyone's marginal decisions — to achieve the desired initial distribution, after which competitive trade does the rest. This requires that the planner:

- a. Can observe each individual's true endowment and welfare capacity.
- b. Can impose transfers that are truly non-distortionary.

These assumptions are absurd once this is made explicit.

27. All the assumptions are restrictive and never achieved in the real world. Essentially the first theorem requires that markets are complete<sup>20</sup> and competitive (a circular definition in itself) and the second theorem describes a first-best world that does not exist.

28. Furthermore, the Pareto criterion embedded in neoclassical welfare theory is silent on distribution. There is no reason to suppose that the "best" Pareto efficient point will be selected by the market without intervention — only that some Pareto efficient point will be reached. An outcome in which one person owns everything and everyone else owns nothing can be Pareto efficient. The theory delivers no guarantee that the consumer surplus is distributed equitably.

29. Pareto efficient outcomes also rely on equilibrium theory and sloping supply and demand curves, which as MGUG outlined in its main submission also fail the empirical test.

30. The notion of consumer welfare was originally anchored in neoclassical economics with Alfred Marshall in the 19<sup>th</sup> century, who introduced the concept of consumer surplus— the gap between what consumers would be willing to pay and what they actually pay. Maximising this

<sup>19</sup> Notable contrasts that outperform neoliberalism are China/ Korea/ Taiwan which outgrew liberal peers during state-led phases, China, Bolivia, Costa Rica in poverty reduction via direct state redistribution (vs trickle down), Taiwan/ South Korea, Nordic use of public R&D funding to accelerate innovation

<sup>20</sup> There must be markets for every good, every contingency, and every time period — including futures markets and contingent claims markets for all uncertain states of the world. In practice, markets for many risk-contingencies and future goods simply do not exist (Arrow-Debreu completeness is never achieved).

surplus in **competitive equilibrium** became the standard against which market outcomes were measured.

31. In fact, much of neoclassical theory assume equilibrium outcomes, including in macroeconomic theory (basis of DSGE and GE models). This assumption is entirely incompatible with the nature of the economic system which is complex adaptive (CAS) and evolutionary in nature as discussed above.
32. The neoclassical case for markets is internally coherent but depends on a set of conditions — atomistic competition, free entry, perfect information, no externalities, rational agents — that are simultaneously assumed and routinely violated.
33. **In summary the implied objective in s52A that social welfare is maximised by competitive markets is not empirically proven in the real world.**

### NPV=0 Equivalence to Social Welfare Maximisation Formula

34. The Commission, as it appears, steers and judges its settings on whether the outcome *for suppliers* across a regulatory period demonstrates NPV=0. This is a quantifiable expression of the competitive market outcome statement that is equated with social welfare maximisation.
35. The theoretical chain from NPV=0 to social welfare maximisation can be examined and then systematically identified where empirical tests expose the fractures.

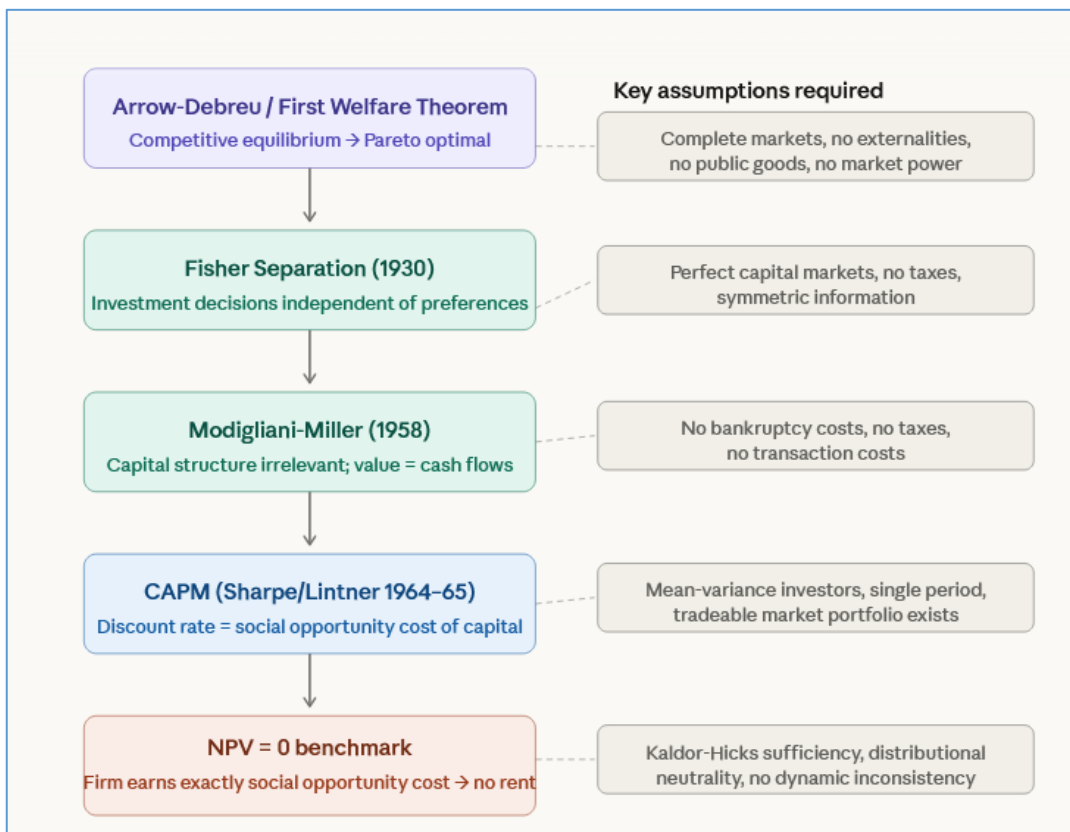
#### The building blocks

36. The argument that NPV=0 maximises social welfare rests on a multi-step derivation rooted in the Arrow-Debreu framework and its successors.
  - a. **Step 1 — Pareto efficiency and competitive equilibrium.** The First Welfare Theorem (Arrow 1951; Debreu 1959) establishes that a competitive equilibrium is Pareto optimal: no one can be made better off without making someone else worse off. This is the foundational linkage between market outcomes and welfare.
  - b. **Step 2 — The Fisher Separation Theorem (1930).** Irving Fisher showed that, in *perfect* capital markets, a firm's investment decisions are independent of shareholders' preferences about consumption timing. Managers need only maximise NPV; shareholders can then use capital markets to rearrange their consumption profiles to match their preferences. This separates the firm's real investment problem from individuals' intertemporal consumption problem.
  - c. **Step 3 — Modigliani-Miller (1958, 1963).** In frictionless markets (*no taxes, no bankruptcy costs, symmetric information*), capital structure is irrelevant to firm value. This means the discount rate used in NPV calculations is just the opportunity cost of capital — what investors could earn elsewhere at equivalent risk — not distorted by financing choices.
  - d. **Step 4 — The CAPM discount rate as a shadow price.** Under CAPM (Sharpe 1964; Lintner 1965; Mossin 1966), the required return on a project equals the risk-free rate plus a risk premium proportional to the project's contribution to market-wide (non-diversifiable) risk. This  $\beta$ -scaled rate is interpreted as a social opportunity cost: it represents what the economy foregoes by committing capital to this project rather than the market portfolio. **When a project earns exactly this rate, NPV = 0, meaning it just covers the social opportunity cost of capital.**

- e. **Step 5 — NPV = 0 as a welfare-neutral threshold.** Projects with  $NPV > 0$  generate returns above the social opportunity cost, creating surplus that can (in principle) compensate losers and leave everyone better off — a potential Pareto improvement (Kaldor-Hicks criterion). Projects with  $NPV = 0$  are exactly break-even: they neither destroy nor create surplus relative to the benchmark. **A competitive firm forced to  $NPV = 0$  (e.g., by regulatory price control) is therefore, under these conditions, consuming no rent — which is the social welfare ideal for a regulated monopoly.**

37. This has a regulatory corollary. In NZ Commerce Act Part 4 terms, this logic underlies WACC-based price control: if a regulated firm earns revenues with a present value equal to the present value of its efficient costs (including a WACC-based return),  $NPV$  of the regulatory contract = 0, and no excess rent is extracted from consumers. This is the BBM/RAB framework's normative foundation.

38. This chain is visualised in *Figure 2*



*Figure 2: Theoretical Chain to maximising social welfare*

39. It is important to note that in *Figure 2* each arrow shown requires the full stack of assumptions above it **to hold simultaneously**.

#### Where the theoretical assumptions fail empirically

40. The literature identifies at least seven distinct fault lines, each corresponding to a broken assumption in the chain above. MGUG has commented on some of these already in its submissions:

41. **1. Market incompleteness and missing prices** (Arrow-Debreu failure)

The First Welfare Theorem requires complete contingent-claims markets — a price for every state of the world. Empirically this is severely violated. Arrow (1964) himself acknowledged that insurance markets for many risks do not exist. Environmental externalities (CO<sub>2</sub>, habitat destruction) have no market price. Future generations' preferences are structurally unrepresented in current capital markets. *This means the discount rate used in NPV calculations does not capture full social cost* — a problem extensively documented in the climate literature (Stern 2006; Nordhaus 2007) where the social cost of carbon cannot be inferred from observed market prices. In NZ infrastructure contexts, the social value of gas network optionality for dry-year electricity firming is similarly unpriced in any observable market.

42. **2. The CAPM is empirically rejected as a cross-sectional pricing model**

The CAPM's empirical record is poor. Fama and French (1992) showed that  $\beta$  alone does not explain cross-sectional variation in stock returns; size and book-to-market factors add substantial explanatory power. Roll (1977) demonstrated that the CAPM is untestable in principle, because the true market portfolio is unobservable and every test is a joint test of the model and the proxy. Jagannathan and Wang (1996) show that conditional CAPM variants perform better but still leave substantial anomalies. *For regulators, this is critical: the  $\beta$  estimates used to set the WACC (and thus the NPV=0 anchor for regulated prices) rest on a model that is empirically rejected in its unconditional form.* The Commerce Commission's own expert witnesses acknowledge the imprecision of  $\beta$  estimation for regulated utilities — comparator firms are internationally sourced, leverage-adjusted, and selected with significant discretionary judgment (see Commerce Commission 2023 DPP3 WACC decisions).

43. **3. Modigliani-Miller violations: taxes, distress costs, and the leverage adjustment problem**

MM holds in a world without taxes or bankruptcy costs. Empirically both matter substantially. The tax shield on debt means that a leveraged firm has higher value than an unlevered one (MM Proposition I with taxes, 1963). Real-world capital structures are influenced by tax incentives, agency costs, and information asymmetries (Myers and Majluf 1984; pecking order theory). For regulatory WACC, this creates a practical problem: the re-levering of asset betas from comparator firms to the regulated entity's *assumed capital structure* depends on the MM leverage formula, but that formula's empirical validity is contested. Small errors in the leverage adjustment compound into material differences in the allowed return, and hence in whether the regulatory NPV=0 target is achievable.

44. **4. Information asymmetry and the efficiency of market prices**

The Fisher Separation Theorem assumes symmetric information. Grossman and Stiglitz (1980) identified the information paradox: if prices perfectly reflect all information, there is no incentive to acquire information, so prices cannot in fact fully reflect it. Empirical evidence of persistent return anomalies — momentum (Jegadeesh and Titman 1993), post-earnings announcement drift, IPO underpricing — confirms that prices do not fully impound all available information. For NPV analysis, this matters because the discount rate is extracted from capital market prices; if those prices are systematically biased, the inferred opportunity cost of capital is wrong.

45. **5. Kaldor-Hicks and the neglect of distribution**

Even where  $NPV > 0$  projects generate aggregate surplus, the Kaldor-Hicks criterion merely requires that winners *could* compensate losers — *not that they do*. Arrow and others (Scitovsky 1941; Blackorby and Donaldson 1990) showed that potential Pareto improvements are neither

transitive nor sufficient for welfare improvement without an explicit social welfare function that specifies distributional weights. In regulatory contexts, this is directly relevant: **NPV=0 pricing that is welfare-neutral at the aggregate level may impose significant distributional harm (e.g., high fixed connection charges on low-income households), which conventional NPV analysis ignores entirely.**

**46. 6. Time-inconsistency and the regulatory commitment problem**

Dynamic models (Kydland and Prescott 1977; Helm, Jenkinson, and Mayer 1988) demonstrate that governments and regulators face time-inconsistency problems: an optimal ex-ante pricing policy (allowing NPV=0 returns to incentivise investment) becomes suboptimal ex-post once the investment is sunk. A regulator can maximise short-run consumer welfare by cutting prices to marginal cost after asset commitment. Investors anticipating this will under-invest. This means that even if all the static NPV=0 welfare conditions held, the dynamic game between regulator and regulated firm systematically distorts the social optimum — the Averch-Johnson effect (1962) being one well-known static variant. Empirically, studies of UK utility regulation post-privatisation (Newbery 1999) document persistent under-investment attributable to regulatory risk.

**47. 7. Behavioural failures and managerial deviation from NPV maximisation**

Even accepting all the theoretical machinery, there is extensive empirical evidence that managers do not in fact maximise NPV. Jensen and Meckling (1976) document agency costs; empire-building managers over-invest. Shleifer and Vishny (1997) survey limits to arbitrage that prevent markets from correcting these deviations. Malmendier and Tate (2005) show that CEO overconfidence leads to systematic NPV-negative acquisitions. Roll (1986) proposed the hubris hypothesis for mergers. These behavioural and agency failures mean that even a theoretically valid NPV=0 benchmark cannot be assumed to be implemented in practice by private actors operating without tight regulatory constraints.

**48. The assumptions and empirical violations are summarised in *Table 2***

*Table 2: Empirical violation of theoretical chain*

<b>Assumption</b>	<b>Theory it supports</b>	<b>Empirical status</b>
Complete contingent markets	First Welfare Theorem	Strongly violated; externalities pervasive
CAPM holds unconditionally	WACC = social opportunity cost	Empirically rejected (Fama-French 1992; Roll 1977)
No taxes or distress costs	MM capital structure irrelevance	Violated; tax shields and bankruptcy costs material
Symmetric information	Fisher Separation, price efficiency	Violated; anomalies persist (Grossman-Stiglitz 1980)
Compensation actually paid	Kaldor-Hicks → welfare improvement	Hypothetical only; distribution ignored
Regulatory commitment credible	Dynamic efficiency of NPV=0	Time-inconsistency endemic (Kydland-Prescott 1977)

Assumption	Theory it supports	Empirical status
Managers maximise NPV	Firm behaviour tracks theory	Violated by agency costs and behavioural biases

### The reality of multi-business firms

49. A key point touched on by MGUG is that multi-business firms like Vector, Powerco, and Firstgas do not behave according to a hypothetical single business, and approximating them all to this theoretical construct can also undermine social welfare outcomes.
50. The building block/RAB/WACC framework rests on a clean theoretical foundation: a single-purpose regulated firm whose entire asset base serves a defined regulated function, whose cost of capital reflects only the risks of that regulated activity, and whose revenues need only cover the efficient costs of that one activity. The NPV=0 principle — that allowed revenues should exactly compensate investors for their capital over time — only holds under these conditions. Reality systematically violates each of them.

### The WACC Is Wrong When the Firm Has Multiple Risk Profiles

51. The most fundamental problem is that WACC is estimated for a benchmark firm, typically using comparator companies. But when the regulated entity also operates in unregulated markets, its observed equity beta and capital structure reflect a blend of risks — some regulated, some not.
52. The direction of bias cuts both ways depending on circumstances. If the unregulated businesses are riskier than the regulated ones (which is common — competitive markets carry more systematic risk than regulated monopoly returns), then using the observed or comparator-derived WACC overstates the true risk of the regulated assets, allowing an excessive return to be charged to regulated consumers. Conversely, if the unregulated businesses are more stable or counter-cyclical, the WACC may be understated, potentially deterring investment in the regulated network.
53. Regulators, including the Commission typically respond by using a "notional" or "benchmark" standalone firm WACC — but this creates its own problem: the benchmark firm does not exist. **The actual firm's financing decisions, tax position, and investor expectations are shaped by its whole portfolio.** Applying a hypothetical standalone WACC to an actual diversified firm severs the link between the regulatory model and the firm's real cost of capital, which is the very thing the WACC is supposed to capture.

### Asset Base Contamination and Cost Allocation

54. The RAB is supposed to contain only assets used in providing the regulated service. But within a firm that operates across multiple businesses, many assets are genuinely shared — head office functions, IT systems, treasury operations, land, brand, management capacity. The allocation of shared assets to the RAB, and therefore the return earned on them, is inherently arbitrary.
55. This creates powerful incentive problems that regulators have consistently struggled to contain:
  - a. **Gold-plating across business lines.** If a shared asset is allocated partly to the RAB, the firm has an incentive to over-invest in it, because the regulated business earns a guaranteed return on RAB assets while the unregulated business benefits from the same asset at effectively subsidised cost.
  - b. **Transfer pricing and cost shifting.** Firms can shift costs from unregulated businesses (where they bear full commercial risk) into regulated cost allowances (where they are

recovered from consumers with certainty). Operating expenditure is particularly vulnerable: internal service charges, management fees, and overhead allocations can be structured to load costs onto the regulated entity.

- c. **Regulatory asset base inflation.** Assets that serve both regulated and unregulated functions may be capitalised into the RAB at their full value even though a significant portion of their economic benefit flows to the unregulated business.

#### The Financeability Problem Is Obscured

56. The building block model assumes that the regulated entity needs to finance its RAB — that is, the  $WACC \times RAB$  calculation represents a real cost of raising capital for that specific activity. But a diversified firm **raises capital at the group level**, not the subsidiary level. Its actual cost of debt is determined by the consolidated balance sheet, the diversity of its cash flows, and the implicit or explicit support of its parent or affiliated businesses. This means:
  - a. The regulated business may benefit from **cross-subsidised borrowing** — cheaper debt because lenders take comfort from the group's non-regulated cash flows — but the regulatory model still allows the full standalone WACC. The firm earns a spread between its actual financing cost and the allowed cost<sup>21</sup>.
  - b. Conversely, if the unregulated businesses are distressed, the regulated entity's financeability may be genuinely impaired even if its own operations are sound — as some water companies in England demonstrated when private equity ownership structures layered corporate debt above otherwise stable regulatory assets.
  - c. The "financeability" tests regulators apply are entirely based on the notional standalone firm, making them poor predictors of the actual firm's financial health or investment

#### The Tax Allowance Is Systematically Distorted

57. Regulatory models typically include a tax allowance in the building blocks, calculated on the basis of the regulated business as if it were a standalone taxpayer. But diversified groups pay tax on consolidated profits. Losses or tax shields in unregulated businesses can offset taxable income that the regulated business generates, meaning the firm pays less tax than the regulatory model assumes — but still recovers the full modelled tax allowance from consumers.
58. This is not a marginal problem. Large diversified infrastructure groups routinely use depreciation in unregulated businesses, debt interest deductions at the group level, or inter-company financing structures to minimise consolidated tax — while continuing to receive the regulatory tax allowance that was designed for a fully taxed standalone entity.

#### Economies of Scope Are Not Passed Through to Consumers

59. The standalone firm model assumes that the regulated firm's costs are those of a firm doing only the regulated activity. A multi-business firm, however, realises economies of scope — shared procurement, shared workforce skills, shared infrastructure, shared corporate governance. These genuine cost savings accrue to the diversified firm but are invisible to the regulator benchmarking against hypothetical standalone operators.
60. The regulatory settlement is therefore systematically *more generous* than necessary: the allowed revenues are calibrated to what a standalone firm would need, but the actual firm's costs are lower due to scope economies. The surplus is retained by shareholders rather than passed to consumers through lower prices.

---

<sup>21</sup> An example noted by MGUG in relation to Powerco

### The Regulated Business as an Internal Capital Source

61. A regulated business with a stable, guaranteed revenue stream and a defined RAB provides something extremely valuable to a diversified group: predictable, bondable cash flows that can support borrowing for other businesses. The regulated entity effectively functions as an internal bank or anchor for the group's balance sheet.
62. This undermines the WACC framework in two ways. First, ***the regulated business's contribution to group value is greater than its standalone NPV because of this option value*** — yet the regulatory model only compensates for standalone economics. Second, groups have incentives to extract value from the regulated entity (through dividends, intercompany loans, or special dividends funded by debt) in ways that do not show up in the regulatory cost model but impair the long-term investment capacity of the network.
63. The UK water sector has provided the most visible examples of this dynamic, where regulatory ringfencing provisions have been repeatedly tested by private equity owners using the RAB as collateral for group-level debt.

### Information Asymmetry Is Amplified

64. The entire regulatory model depends on the regulator having sufficient information to set efficient cost allowances. Information asymmetry between the regulator and the regulated firm is already the central challenge in incentive regulation theory (the Laffont-Tirole framework). When the firm operates across multiple businesses, this asymmetry becomes far more severe:
  - a. The firm can obscure the true cost of the regulated business within consolidated accounts.
  - b. Benchmarking against comparable firms becomes harder when no true comparator exists.
  - c. Internal transfer prices for shared services are unobservable to the regulator and subject to strategic manipulation.
  - d. The firm's overall profitability provides little signal about regulated-business efficiency when unregulated businesses dominate group returns.
65. Regulators are therefore setting prices with less reliable information than the theoretical framework requires.

### Incentive Compatibility Breaks Down

66. CPI-X / building block regulation is designed to give the firm an incentive to outperform cost benchmarks — efficiency gains within the regulatory period are retained by shareholders, creating a profit motive for cost reduction. But in a diversified firm, this incentive is diluted and distorted:
  - a. Management attention and capital allocation are spread across multiple businesses; the regulated entity may not receive optimal managerial focus.
  - b. Cost outperformance in the regulated business may be achieved by loading costs onto unregulated entities — an "efficiency gain" that is illusory at the group level.
  - c. The firm may strategically underperform in the regulated business near the end of a regulatory period to preserve high cost benchmarks for the next reset, while using profits from unregulated businesses to sustain overall group returns — making it harder for regulators to detect the sandbagging.

### The Ownership Separation Problem

67. In several jurisdictions, including New Zealand the same corporate group holds regulated assets in multiple regulated industries — for instance, owning both electricity distribution and gas distribution networks. Each is regulated separately, without consideration of portfolio effects:
- a. Structuring financing to maximise the combined regulatory tax shield.
  - b. Allocating shared costs to the regulated entity with the most generous cost pass-through provisions.
  - c. Timing capital expenditure across businesses to exploit differences in regulatory period timing.
  - d. Using WACC differentials between regulated businesses to shift internal financing costs to the entity with the higher allowed return.
  - e. Amplifying narratives across the different regulated industries. For example, the narrative of gas demand collapse is used to justify both accelerated depreciation in gas networks and increased CAPEX allowances for electricity networks burdening the consumer with two high price rises when the firm has the ability (and often belief) to mitigate the narrative it presents to the regulator<sup>22</sup>.

### The Fundamental Tension

68. All of these problems flow from a single underlying tension: the building block / RAB / WACC framework is a *firm-level model* designed for a *market-level problem*. It asks what revenues a specific firm needs, but it answers using theory that assumes a single-purpose firm in a single market. When the actual firm is a diversified enterprise with regulated and unregulated subsidiaries, shared assets, consolidated financing, and group-level tax management, the model's outputs are an approximation at best and systematically biased at worst.
69. Regulators have responded with partial fixes — ringfencing requirements, consolidated accounts reporting, cost allocation rules, related-party transaction controls — but none of these fully resolve the fundamental mismatch. They are patches on a framework whose theoretical foundations were not designed for the institutional reality of modern infrastructure conglomerates.
70. The deepest implication is that the NPV=0 principle — the theoretical guarantee that investors receive fair compensation and no more — cannot be reliably maintained when the "firm" being regulated is actually a node within a larger corporate structure whose economics the regulator cannot fully observe or control.
71. This cross-industry portfolio effect has no place in the single-market theoretical framework but is a routine feature of large infrastructure conglomerates.

---

<sup>22</sup> This is the example that MGUG used to explain Vector's apparent behaviour and motives in undermining its own gas business.

## Implications for NZ regulatory practice

72. It should be sufficient to show the naivety of the assumptions carried in the regulatory framework. Nothing that the Commission uses in its IM rests on any empirically solid ground. Distributional effects are addressed only obliquely through affordability provisions, not through formal welfare weights. The failure to even consider welfare impacts beyond “affordability” for a hypothetical representative agent (the “average household”) is a major problem for the economy as MGUG has continually tried to demonstrate.
73. There are however practical tests that the Commission can both adopt and apply to determine whether its current draft decision improves or worsens welfare outcomes.

## The Problem with the Representative Consumer Fiction

74. The standard regulatory welfare test — does the price control deliver efficient prices that a hypothetical welfare-maximising consumer would prefer? — abstracts away from the actual distribution of impacts across real consumers. It treats the regulated market as if it serves a single average household with average income, average consumption, average access to alternatives, and average ability to respond to price changes. This fiction is analytically convenient but empirically false, and it means regulators can declare a price control “welfare-improving” while simultaneously causing serious harm to identifiable groups of actual consumers.
75. The challenge is methodological and institutional as well as technical. Testing actual welfare outcomes requires regulators to move from ex ante modelling of hypothetical impacts to ex post measurement of realised impacts on differentiated populations — which demands different data, different skills, and a different relationship with the consumers being regulated.

## Distributional Impact Analysis Across Consumer Segments

76. The most immediate improvement over the representative consumer is systematic segmentation of the consumer population and measurement of price control impacts across those segments.
77. The analytical framework would involve identifying the relevant dimensions of heterogeneity: income deciles, geographic location (urban vs rural, network periphery vs core), household composition (single adults, families with children, elderly households), tenure type (owner-occupiers vs renters, where the person paying the bill may differ from the person bearing the welfare impact), business size, and disability status where service dependency is higher.
78. For each segment, the Commission would model and then measure the share of household income or business cost absorbed by the regulated price, the responsiveness of consumption to price changes (low-income households typically have much lower price elasticity for essential services because they have already minimised consumption), and the availability of substitutes.
79. This goes beyond what most regulators currently do. The UK water sector's affordability metrics, for instance, track the percentage of households spending more than a defined threshold of income on water bills — but this is typically modelled prospectively using survey data rather than measured from actual billing records matched to income data.
80. The data infrastructure required is significant: matching regulated billing data with administrative income data (tax records, benefits data) at the household level, ideally through secure data linkage rather than relying on self-reported survey income. Several jurisdictions have the administrative architecture to do this — New Zealand's Integrated Data Infrastructure

(IDI) is a world-leading example of linking administrative datasets — but regulators have rarely exploited such resources systematically.

#### Expenditure Share and Affordability Tracking

81. A direct welfare measure for essential services is the expenditure share — what proportion of disposable income does a household actually spend on the regulated service? Unlike a representative consumer average, tracking this across the income distribution reveals the regressive or progressive incidence of regulatory decisions.
82. A price control that looks efficient in aggregate — minimising deadweight loss relative to a competitive benchmark — can simultaneously be highly regressive if network costs are socialised through uniform tariffs while consumption benefits are concentrated among wealthier, higher-consuming households<sup>23</sup>.
83. Tracking this over time across regulatory periods would allow regulators to test whether successive price controls have improved or worsened the distributional incidence of regulated prices — something entirely invisible in representative consumer welfare assessments.<sup>24</sup>
84. The regulator could establish explicit affordability thresholds — not just as policy aspirations but as binding outcome measures that the price control is required to meet or improve. If a threshold proportion of low-income households must spend no more than a defined share of income on the regulated service, this becomes a welfare constraint on the optimisation, not just a consideration.

#### Service Quality Differentiation and the Non-Price Welfare Dimensions

85. Consumer welfare in regulated industries is not purely a function of price. Service quality, reliability, accessibility, and continuity are often more important welfare dimensions for vulnerable consumers than marginal price changes — and they are also more easily manipulated by regulated firms under revenue-cap regulation, which creates incentives to maintain revenue by cutting quality where possible.
86. Measuring actual quality outcomes by consumer segment rather than system-wide averages is essential. A regulated electricity network may report a system average interruption duration index that looks acceptable, but if outages are highly concentrated in older, lower-income, rural, or Māori communities, the average is concealing serious distributional harm.
87. The welfare test should therefore include:
  - a. Interruption frequency and duration by geographic and demographic segment, not just network-wide averages.
  - b. Response times for fault repair across customer types.
  - c. Accessibility of customer service for consumers with disabilities, language barriers, or low digital literacy.
  - d. Connection costs and timeframes for new customers, which disproportionately affect rural communities and new housing developments in lower-income areas.

---

<sup>23</sup> A definite feature of GDB pricing methodology, especially Vector's current pricing for households where the pricing structure is 100% fixed cost, favouring larger, wealthier households against low income, low user households.

<sup>24</sup> Sufficient historical data already exists from published pricing since DPP1 to test this.

88. This moves the quality dimension of regulation from an input (what quality standards does the firm commit to meeting?) to an outcome (what quality did different groups of actual consumers actually experience?).

#### Revealed Preference and Demand-Side Welfare Measurement

89. Rather than modelling what consumers *should* prefer based on assumed preferences, regulators could measure what consumers *actually do* in response to regulatory changes — revealed preference approaches.
90. Consumption response tracking after price changes provides direct evidence of welfare impacts. If low-income households reduce consumption of an essential service following a price increase, this is direct evidence of welfare harm — households are being pushed below their preferred consumption level. Tracking this through billing data disaggregated by customer type, before and after a price control reset, provides a concrete welfare test.
91. For services with some degree of consumer choice or switching (where applicable), tracking switching behaviour, complaint rates, and hardship scheme uptake provides additional revealed preference signals. A spike in payment difficulties following a price control reset is direct evidence that the price path has exceeded the welfare tolerance of a portion of the consumer population — a signal entirely absent from the representative consumer framework.

#### Hardship and Debt Metrics as Welfare Indicators

92. For essential services with inelastic demand, price increases that exceed household budgets manifest not in reduced consumption but in non-payment, debt accumulation, and hardship programme uptake. These are highly direct welfare harm indicators.
93. Tracking systematically:
  - a. Rates of disconnection or service limitation by income band and geographic area.
  - b. Growth in regulated-service debt as a share of household income across deciles.
  - c. Uptake and adequacy of hardship fund provision by regulated firms.
  - d. Rates of self-disconnection (where metering allows it) or rationing behaviour such as reduced heating in winter.
94. These metrics provide a direct empirical test of whether the price control has kept the service genuinely accessible to the full consumer population or has effectively priced out marginal consumers — something the representative consumer model cannot detect by construction.
95. The Commission could require regulated firms to publish these metrics annually as part of their information disclosure obligations, creating a longitudinal record against which the welfare impacts of successive regulatory periods can be assessed.

#### Geographic and Spatial Welfare Testing

96. Regulated network industries are inherently spatial — the cost of serving a consumer depends critically on location, and uniform national or regional tariffs socialise these cost differences across the consumer base. This creates systematic cross-subsidies that have welfare implications for both those who pay above their cost-reflective share and those who are subsidised. Spatial welfare analysis would map the relationship between:
  - a. Actual cost to serve by geographic unit (which the regulated firm knows but rarely publishes).
  - b. Prices charged by geographic unit (typically uniform within a regulatory area).

- c. Consumer income and vulnerability by geographic unit (from census and administrative data).
97. This would reveal, for instance, whether the cross-subsidy flowing from urban to rural consumers is progressive (rural consumers are on average lower income) or regressive (rural areas include both wealthy lifestyle properties and remote communities with no alternatives). The welfare implication differs dramatically between these cases, and the representative consumer framework cannot distinguish them.

#### Long-Run Investment and Intergenerational Welfare

98. Consumer welfare in capital-intensive network industries has a strong intergenerational dimension that the representative consumer framework typically collapses into a single period. The NPV=0 principle ensures investors are fairly compensated, but it says nothing about the timing of consumer payments or the distribution of costs across current and future consumers.
99. Testing intergenerational welfare would involve:
- a. Tracking the ratio of capital expenditure to depreciation over time — persistent underspending relative to depreciation is a signal that current consumers are benefiting at the expense of future consumers who will face either a deteriorating network or a future investment catch-up. Conversely an accelerated depreciation profile advantages future consumers at the expense of current consumers.
  - b. Measuring actual asset condition and age profiles against the depreciation assumptions built into the RAB, to test whether the regulatory settlement is preserving the real capital base.
  - c. Comparing actual service reliability trends with what the investment profile would predict, as a forward-looking test of whether the price control is funding adequate maintenance of the network.
100. This matters particularly because regulated firms under revenue caps have strong incentives to defer maintenance and capital expenditure — costs that fall within the regulatory period but whose consequences (network deterioration, future reliability failures) arrive in future periods when a different set of consumers pays the price.

#### Comparative and Benchmarking Approaches

101. Where multiple regulated entities exist within a jurisdiction — as with electricity distribution businesses in New Zealand or Australia — the Commission has the opportunity to test welfare outcomes comparatively rather than just absolutely.
102. Benchmarking welfare outcomes across comparable regulated firms, controlling for network characteristics and consumer demographics, can reveal whether some firms are delivering systematically better welfare outcomes than the regulatory settlement requires, suggesting either that the price control is more generous than necessary or that some firms are outperforming efficiency expectations. Conversely, firms that consistently underperform on both price and quality dimensions relative to peers are a signal that the regulatory framework is not delivering welfare outcomes even by its own standards.
103. This comparative approach is distinct from the efficiency benchmarking regulators already do on costs — it would benchmark outcomes on the consumer side, including price levels, quality, affordability, accessibility, and hardship indicators.

### Independent Consumer Welfare Audits

104. Beyond internal regulatory measurement, an institutional mechanism worth considering is periodic independent consumer welfare audits — conducted not by the Commission (which has an inherent interest in validating its own decisions) and not by the regulated firm (obvious conflict of interest) but by an independent body with access to the full data set.
105. Such audits would assess whether the price control as implemented has delivered the welfare outcomes the regulatory decision projected, identify which consumer segments have experienced welfare gains or losses, examine the distribution of actual returns to investors relative to what the WACC calculation allowed, and make findings publicly available to inform the next regulatory period.
106. This creates an accountability mechanism that is entirely absent from current regulatory frameworks, most of which assess the quality of the process of price setting (consultation, methodology transparency, legal compliance) without ever measuring whether the outcome served consumer welfare as promised.

### The Institutional Challenge

107. All of these approaches face a common institutional obstacle: they require the Commission to be accountable for outcomes rather than just process quality, which is a fundamentally more demanding standard. Current regulatory frameworks in most jurisdictions evaluate regulators on whether they followed correct procedures, consulted appropriately, applied stated methodologies consistently, and made legally defensible decisions. None of this requires the Commission to demonstrate that actual consumers were made better off.
108. Shifting to outcome-based welfare testing would require not just new data and analytical methods but a different conception of regulatory accountability — one that accepts that a technically correct price control process can still produce harmful welfare outcomes, and that the regulator has a responsibility to detect and respond to those outcomes rather than simply defending the integrity of the process that produced them.
109. This is ultimately a governance question as much as a technical one, and it is worth noting that it sits in direct tension with one of the other legacies of the neoliberal reform era: the design of regulatory agencies as procedurally accountable but outcome-insulated institutions, shielded from political pressure precisely by being evaluated on process rather than result