

Quantum Monte Carlo Methods for Large-Scale Macroeconomic Models: A Unified Computational Framework for High-Dimensional State Spaces and Stress Testing

Sid Ahmed ZENAGUI

`sidahmed.zenagui@univ-temouchent.edu.dz`

University Of Aln Temouchent <https://orcid.org/0009-0002-5333-5106>

Research Article

Keywords: quantum computing, amplitude estimation, Monte Carlo methods, heterogeneous agent models, deep learning, stress testing, computational economics, macro-financial risk

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11402714/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Quantum Monte Carlo Methods for Large-Scale Macroeconomic Models: A Unified Computational Framework for High-Dimensional State Spaces and Stress Testing

Sid Ahmed Zenagui*

University of Ain Temouchent, Algeria

Laboratory of Markets, Employment, Legislation, and Simulation in the Maghreb Countries

This version: June 2026

Abstract

Large-scale macroeconomic models with rich heterogeneity and aggregate risk confront two intertwined computational bottlenecks: the curse of dimensionality in solving high-dimensional functional equations, and the slow $\mathcal{O}(N^{-1/2})$ convergence of Monte Carlo estimators used in policy evaluation and stress testing. This paper develops and analyses *QMC-Macro*, a unified computational framework that combines a deep-learning surrogate for the model’s policy and value functions with quantum amplitude estimation (QAE) for the expectation functionals that summarise macro-financial risk. We do not claim a new economic model or a first-of-its-kind algorithm; rather, we consolidate, formalise, and extend an existing and rapidly maturing literature—most directly the macroeconomic quantum Monte Carlo programme of [Skavysh et al. \(2023\)](#)—into a single, modular pipeline with explicit error accounting. We provide rigorous results bounding the total estimation error as the sum of a neural-approximation term, a distribution-loading term, and an amplitude-estimation term that scales as $\mathcal{O}(N_{\text{or}}^{-1})$, a quadratic improvement over classical sampling. Using a calibrated income-fluctuation economy solved to a mean $\log_{10}$ Euler-equation residual below -8 , we show that a feed-forward surrogate reproduces the optimal policy with a mean relative error below 0.5% and R^2 above 0.999, and that a faithful classical emulator of iterative QAE attains the theoretical $\mathcal{O}(N_{\text{or}}^{-1})$ scaling, implying sample-complexity speedups that rise from roughly $25\times$ at percent accuracy to several hundredfold in the deep tail. We discuss feasibility on noisy intermediate-scale quantum (NISQ) hardware, the resource and circuit-depth trade-offs, and the boundaries of the quantum advantage. The framework is offered as a refinement and unification of existing tools rather than a replacement for them, and we are explicit about the assumptions under which the advantage materialises.

Keywords: quantum computing; amplitude estimation; Monte Carlo methods; heterogeneous-agent models; deep learning; stress testing; computational economics; macro-financial risk.

JEL classification: C63; C45; C61; E37; G17; C15.

*University of Ain Temouchent, Algeria; Laboratory of Markets, Employment, Legislation, and Simulation in the Maghreb Countries (LAMEL). Email: Sidahmed.zenagui@univ-temouchent.edu.dz. The author thanks, without implicating, seminar participants for helpful comments. All errors are the author’s own.

Contents

1	Introduction	5
2	Related Literature	7
2.1	Solving high-dimensional dynamic models	7
2.2	Deep learning for economics	7
2.3	Quantum algorithms for Monte Carlo	8
2.4	Quantum computing and the macro-financial system	8
3	Background: Quantum Computation for Economists	8
3.1	Qubits, states, and measurement	9
3.2	Gates and circuits	9
3.3	Encoding an expectation as an amplitude	9
3.4	Grover iteration and amplitude amplification	9
3.5	Canonical, maximum-likelihood, and iterative amplitude estimation	10
3.6	The NISQ reality	10
4	The Macroeconomic Environment	10
4.1	Households	10
4.2	Optimality and the recurring expectation	11
4.3	Where the dimensionality comes from	11
4.4	Stationary equilibrium and the cross-sectional distribution	12
4.5	The aggregate-risk extension	12
4.6	From the model to a loss functional	13
5	The Computational Problem and Classical Benchmarks	13
5.1	Expectation functionals	13
5.2	Classical Monte Carlo	13
5.3	The cost of constructing the integrand	14
5.4	A complexity map	14
6	The QMC-Macro Framework	14
6.1	Stage 1: the model and its ground-truth solution	15
6.2	Stage 2: the deep-learning surrogate	16
6.3	Stage 3: loading the distribution and integrand	16
6.4	Stage 4: amplitude estimation	17
6.5	Stage 5: read-out and risk measures	17
6.6	The comparator and the objective rotation	17
6.7	Qubit accounting	17
6.8	The IQAE schedule	18

7	Methodology and Implementation	18
7.1	Discretising the productivity process	18
7.2	The endogenous grid method	18
7.3	Validating the solution: Euler residuals	19
7.4	Training the surrogate	19
7.5	Classical Monte Carlo benchmark	20
7.6	The amplitude-estimation emulator	20
7.7	Reproducibility discipline	20
8	Theoretical Analysis	20
8.1	Classical and quantum estimation errors	21
8.2	Surrogate-approximation error	21
8.3	Distribution-loading error	22
8.4	The end-to-end bound	22
8.5	Resource complexity	23
9	Data and Calibration	23
9.1	Calibration	23
9.2	Macro-financial data for scenario calibration	24
10	Numerical Results	24
10.1	Solution and surrogate accuracy	24
10.2	Convergence of the estimation stage	25
10.3	Sample-complexity speedups	26
10.4	Stress scenarios	27
10.5	Resource accounting	28
10.6	Solver cost across resolutions	29
10.7	A worked example	29
11	Robustness	30
11.1	Sensitivity to risk aversion	30
12	Extension to HANK and Financial-Sector Models	31
12.1	A HANK environment	31
12.2	A bank-solvency stress functional	31
12.3	What changes and what does not	32
13	Discussion	32
13.1	Anticipating reviewer concerns	32
13.2	Limitations	33
13.3	Policy relevance	33
13.4	From NISQ to fault tolerance	33
13.5	Cryptographic considerations	34
13.6	A research agenda	34

14 Comparison with Classical Alternatives	34
14.1 Quasi-Monte Carlo	34
14.2 Importance sampling for rare events	34
14.3 The quantum-annealing route	35
14.4 Summary of the comparison	35
15 Conclusion	35
A Proofs and Additional Derivations	36
A.1 Derivation of the Euler equation	36
A.2 Contraction property of policy iteration	36
A.3 Full proof of the amplitude-estimation geometry	36
A.4 Loading discretisation error versus qubit count	37
A.5 Discontinuous integrands	37
A.6 Optimal split of the error budget	37
A.7 The Rouwenhorst construction	37
A.8 From tail probabilities to expected shortfall	38
A.9 IQAE confidence intervals	38
B Algorithms	38
C Data Appendix	39
D Reproducibility	40
E Glossary of Quantum Terms	40
F Notation	41

1 Introduction

Modern quantitative macroeconomics has moved decisively toward models that take heterogeneity, occasionally binding constraints, and aggregate risk seriously. Heterogeneous-agent New Keynesian (HANK) models (Kaplan et al., 2018), continuous-time mean-field formulations (Achdou et al., 2022), and models with an explicit financial sector (Brunnermeier and Sannikov, 2014) have become standard tools for analysing monetary and fiscal policy, inequality, and financial stability. These advances come at a cost. The state of such a model is, in general, an entire cross-sectional distribution, and the objects of interest—policy functions, value functions, and the conditional expectations that enter equilibrium conditions—live in spaces whose dimensionality grows rapidly with the richness of the model. Two computational obstacles recur. The first is the *curse of dimensionality*: classical grid- and projection-based solution methods require a number of basis functions or collocation nodes that scales exponentially in the dimension of the state, rendering them infeasible beyond a handful of continuous state variables (Judd, 1998). The second is the *slow convergence of Monte Carlo estimation*: the workhorse for evaluating expectations—impulse responses, welfare functionals, value-at-risk, and the tail probabilities at the heart of supervisory stress testing—attains a root-mean-square error that decays only as $\mathcal{O}(N^{-1/2})$ in the number of simulated scenarios N (Glasserman, 2004). Halving the error requires quadrupling the computational budget, a scaling that is especially punishing for rare-event and deep-tail functionals.

A vigorous recent literature has attacked each obstacle separately. On the solution side, deep neural networks have emerged as flexible function approximators that, for the relevant smoothness classes, can represent high-dimensional policy and value functions with a number of parameters that grows only polynomially in the state dimension (Maliar et al., 2021; Azinovic et al., 2022; Han et al., 2021; Fernández-Villaverde et al., 2024). On the estimation side, quantum algorithms based on amplitude estimation promise a near-quadratic speedup for the Monte Carlo evaluation of expectations (Montanaro, 2015; Rebentrost et al., 2018; Woerner and Egger, 2019; Stamatopoulos et al., 2020). The intersection of these two strands—quantum methods deployed specifically for the economics of stress testing and macroeconomic simulation—was opened by Skavysh et al. (2023), who showed how amplitude estimation and macroeconomic deep learning can be combined, and by a growing body of central-bank and policy analysis assessing the prospects of quantum computing for the financial system (Gorbanyov et al., 2021; Bank for International Settlements, 2024; Fernández-Villaverde and Hull, 2023). The present paper builds directly on this foundation.

What this paper does, and does not, claim. We are deliberately modest about novelty. We do not introduce a new macroeconomic model, we do not propose a new quantum primitive, and we do not assert a first-ever application. Each ingredient of our framework—the income-fluctuation model, the endogenous-grid solution method, the neural surrogate, canonical and iterative amplitude estimation—is drawn from the existing literature and cited as such. Our contribution is one of *unification, formalisation, and extension*. Specifically, we (i) assemble the solution and estimation stages into a single modular pipeline, *QMC-Macro*, with clean interfaces between the classical and quantum components; (ii) provide an explicit, end-

to-end error decomposition that makes transparent how the surrogate-approximation error, the distribution-loading error, and the amplitude-estimation error combine, and under what conditions the quantum advantage survives the classical pre-processing it requires; (iii) supply a fully reproducible numerical study, including a faithful classical emulator of iterative amplitude estimation that reproduces the theoretical $\mathcal{O}(N_{\text{or}}^{-1})$ scaling without access to quantum hardware; and (iv) discuss carefully the noisy-intermediate-scale-quantum (NISQ) feasibility, the resource trade-offs, and the regimes in which the advantage is largest or evaporates. Throughout, we treat the framework as a refinement of existing tools, and we foreground the assumptions under which our claims hold.

Contributions, stated precisely. It is worth being exact about the marginal contribution relative to the closest prior work (Skavysht et al., 2023; Fernández-Villaverde et al., 2024; Herman et al., 2023). We claim four things. First, a *modular architecture* with explicit, auditable interfaces between a classical solution stage, a classical surrogate stage, and a quantum estimation stage, such that each can be improved or replaced independently. Second, an *end-to-end error theorem* (Theorem 1) that decomposes the total error into surrogate, loading, and estimation terms and that assembles in one place bounds which are individually known but not previously combined for this workflow. Third, a *regime characterisation* (Corollary 1 and the surrounding remarks) stating precisely when the quantum advantage survives the classical pre-processing it requires, distinguishing carefully between fixed-absolute and fixed-relative accuracy targets—a distinction that materially changes the ranking of scenarios. Fourth, a *fully reproducible benchmark*, including a faithful classical emulator of iterative amplitude estimation, that lets any reader measure the constants and resource trade-offs without quantum hardware. None of these is a new model or a new quantum primitive; the contribution is consolidation, formalisation, and honest delineation of scope.

The computational problem in a sentence. Most quantities that macroeconomists and supervisors care about can be written as an expectation $a = \mathbb{E}[f(X)]$ of a bounded functional f of a (possibly high-dimensional) random vector X whose law is induced by the solved model and a scenario distribution. For the leading stress-testing application, f is an indicator of a loss exceeding a threshold, so that $a = \mathbb{P}(L(X) > \ell)$ is exactly a tail probability. Classical Monte Carlo estimates a with error $\sqrt{a(1-a)/N}$. Amplitude estimation encodes a as the squared amplitude of a marked quantum state and extracts it with error $\mathcal{O}(N_{\text{or}}^{-1})$ in the number N_{or} of applications of a Grover-type operator (Brassard et al., 2002; Grinko et al., 2021). The role of the deep-learning surrogate is to make the construction of the integrand f —which requires repeated evaluation of the model’s policy functions—cheap and differentiable, and to keep the quantum circuit that loads the distribution shallow enough to be plausible on near-term hardware.

Preview of results. We calibrate a standard income-fluctuation economy and solve it by the endogenous grid method to a mean $\log_{10}$ Euler-equation residual of -8.97 (Table 3), establishing a reliable ground truth. A modest feed-forward network (three hidden layers of width 64) reproduces the optimal consumption policy with a mean relative error of 0.43% and a coefficient of determination of 0.9993, confirming that the surrogate stage introduces a controllably small

error. We then study the estimation stage on a battery of stress scenarios calibrated to macro-financial data. A classical emulator of iterative amplitude estimation attains root-mean-square error scaling as $\mathcal{O}(N_{\text{or}}^{-1})$, in contrast to the $\mathcal{O}(N^{-1/2})$ scaling of classical Monte Carlo (Figure 6). The implied sample-complexity speedup grows with the required precision and with the rarity of the event being priced, reaching several hundredfold for the deepest tail scenario we consider (Table 5). We pair these results with an explicit NISQ resource accounting and a frank discussion of the circuit-depth costs that temper the practical advantage today.

Roadmap. Section 2 positions the paper in the literature. Section 4 sets out the macroeconomic environment and the precise sense in which it is high-dimensional. Section 5 states the computational problem and the classical benchmarks. Section 6 develops the QMC-Macro framework, component by component. Section 8 contains the theoretical analysis: the error decomposition, the query-complexity results, and the combined bound. Section 9 describes the data and calibration. Section 10 reports the numerical results, and Section 11 the robustness analysis. Section 13 discusses feasibility, limitations, and reviewer-anticipated concerns. Section 15 concludes. Proofs, algorithms, data details, and additional results are collected in the appendices.

2 Related Literature

This paper sits at the confluence of four literatures: the numerical solution of high-dimensional dynamic economic models, deep learning for economics, quantum algorithms for Monte Carlo integration, and the nascent study of quantum computing for the financial system. We review each in turn and state precisely how our contribution relates to it.

2.1 Solving high-dimensional dynamic models

The canonical heterogeneous-agent models of Aiyagari (1994) and Krusell and Smith (1998) already strain classical solution methods once aggregate risk is added, because the relevant state includes a distribution. Continuous-time formulations (Achdou et al., 2022) and models with a financial sector (Brunnermeier and Sannikov, 2014) sharpen the challenge. The classical toolkit—value-function iteration, projection, and perturbation (Judd, 1998)—suffers from the curse of dimensionality: tensor-product grids require $\mathcal{O}(g^d)$ nodes for g points per dimension and d dimensions. Our framework does not replace these methods; it uses a standard, well-understood solution (the endogenous grid method of Carroll, 2006) to generate a ground truth against which the surrogate is trained and validated, exactly as a practitioner would.

2.2 Deep learning for economics

A fast-growing literature uses neural networks to approximate policy and value functions in high dimensions. Maliar et al. (2021) reformulate dynamic models as nonlinear regressions solved by stochastic gradient descent; Azinovic et al. (2022) introduce deep equilibrium nets; Han et al. (2021) propose DeepHAM, which represents the cross-sectional distribution by learned generalised moments; and Fernández-Villaverde et al. (2024) survey the programme of “taming

the curse of dimensionality” with deep learning, with continuous-time master-equation solvers developed by Gu et al. (2024). The theoretical underpinning is the universal-approximation property (Cybenko, 1989; Hornik et al., 1989) and its high-dimensional refinements via deep solvers of partial differential equations (Han et al., 2018; Sirignano and Spiliopoulos, 2018). The common thread is that, for the smoothness classes relevant to economic policy functions, a network’s parameter count can grow polynomially rather than exponentially in the state dimension, breaking the classical curse of dimensionality at the solution stage. What this literature does not address is the *estimation* stage: once a high-dimensional model is solved, the expectations that summarise its risk implications must still be computed, and there classical Monte Carlo reasserts its $\mathcal{O}(N^{-1/2})$ tax. Our framework treats the deep-learning solution as an input and targets that residual estimation bottleneck. We use a deliberately simple feed-forward surrogate; the point is not to advance the deep-learning frontier but to quantify the error this stage contributes to the end-to-end pipeline and to keep the construction of the quantum integrand cheap and differentiable.

2.3 Quantum algorithms for Monte Carlo

Grover’s search (Grover, 1996) and the amplitude amplification and estimation framework of Brassard et al. (2002) underlie a near-quadratic quantum speedup for estimating expectations (Heinrich, 2002; Montanaro, 2015). In finance, Rebentrost et al. (2018) price derivatives by amplitude estimation; Woerner and Egger (2019) and Stamatopoulos et al. (2020) develop quantum risk analysis and option pricing; Egger et al. (2020) treat credit risk; and An et al. (2021) combine amplitude estimation with multilevel Monte Carlo. Because canonical amplitude estimation relies on quantum phase estimation and hence deep circuits, a line of work develops phase-estimation-free and iterative variants suitable for near-term hardware (Suzuki et al., 2020; Grinko et al., 2021). Chakrabarti et al. (2021) quantify the resource threshold at which quantum advantage in derivative pricing becomes plausible, and Herman et al. (2023) survey the field. Two themes from this literature shape our design. First, the headline quadratic speedup is a statement about *query complexity*—the number of oracle calls—and translates into a wall-clock advantage only once the oracle (the loading circuit) is cheap; the loading bottleneck is the recurring caveat (Herman et al., 2023; Chakrabarti et al., 2021). Second, the choice of estimator matters for feasibility: canonical AE is asymptotically optimal but deep, whereas iterative and maximum-likelihood variants trade a logarithmic number of extra rounds for drastically shallower circuits. We adopt iterative amplitude estimation (Grinko et al., 2021) as our reference estimator precisely because of its bounded per-round circuit depth and its certified confidence intervals, which a supervisor reporting an exceedance probability requires.

2.4 Quantum computing and the macro-financial system

The most direct antecedent of this paper is Skavysh et al. (2023), who study quantum Monte Carlo for economics—stress testing and macroeconomic deep learning—and demonstrate the combination of amplitude estimation with learned model solutions. Policy institutions have begun to assess the implications: the Gorbanyov et al. (2021) and the Bank for International Settlements (2024) weigh the opportunities for stress testing and macroeconomic analysis against

the cryptographic risks, while [Fernández-Villaverde and Hull \(2023\)](#) solve dynamic programs on a quantum annealer and [Hull et al. \(2020\)](#) survey quantum technology for economists. Related work uses quantum optimisation to reduce systemic risk in financial networks ([Aboussalah et al., 2023](#)) and to forecast crashes ([Orús et al., 2019](#)); [Bouland et al. \(2020\)](#) catalogue the prospects and challenges. Against this backdrop, our marginal contribution is the explicit, modular error accounting that ties the solution and estimation stages together and delineates when the advantage survives in a realistic, end-to-end workflow.

3 Background: Quantum Computation for Economists

Because the intended readership is economists rather than quantum physicists, this section collects, in self-contained form, the minimum quantum-computational apparatus needed to read the rest of the paper. None of the material is original; standard references are [Nielsen and Chuang \(2010\)](#) for the fundamentals and [Herman et al. \(2023\)](#) for the finance-facing summary. Readers familiar with amplitude estimation may proceed directly to Section 4.

3.1 Qubits, states, and measurement

A single quantum bit, or qubit, is a unit vector in $\mathbb{C}^2$ written in Dirac notation as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$. The numbers α, β are *amplitudes*; measurement in the computational basis returns 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$, after which the state collapses to the observed basis vector. A register of n qubits lives in the 2^n -dimensional tensor-product space $(\mathbb{C}^2)^{\otimes n}$, and a general state is a superposition $\sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$ over the 2^n computational basis states, with $\sum_x |\alpha_x|^2 = 1$. The exponential dimension of this space relative to the linear qubit count is the source of the potential computational leverage, and the constraint that information is extracted only through probabilistic measurement is the source of the difficulty: one cannot simply read out all 2^n amplitudes.

3.2 Gates and circuits

Computation proceeds by applying unitary operators (gates) to the register. Single-qubit gates include the Pauli operators and the Hadamard gate H , which maps $|0\rangle \mapsto (|0\rangle + |1\rangle)/\sqrt{2}$ and so creates superposition; two-qubit entangling gates such as the controlled-NOT generate correlations with no classical analogue. A quantum *circuit* is a finite sequence of gates; its *depth* is the number of sequential gate layers, and depth is the binding resource on near-term hardware because physical qubits decohere. The quantity we track throughout—the number N_{or} of applications of the Grover operator—is the dominant contributor to circuit depth in amplitude estimation, which is why the resource accounting of Section 10 matters as much as the speedup.

3.3 Encoding an expectation as an amplitude

The bridge from economics to quantum computation is the encoding of an expectation as a squared amplitude. Suppose we can build a state-preparation unitary P that loads a discrete

distribution $\{p_x\}$ over n qubits, $P|0\rangle_n = \sum_x \sqrt{p_x} |x\rangle$, and a rotation R that, controlled on the value x , rotates an ancilla objective qubit by an angle encoding $f(x) \in [0, 1]$, so that $R|x\rangle|0\rangle = |x\rangle(\sqrt{1-f(x)}|0\rangle + \sqrt{f(x)}|1\rangle)$. The composite $\mathcal{A} = R(P \otimes I)$ then prepares precisely the state (12), in which the probability of measuring the objective qubit in $|1\rangle$ equals $\sum_x p_x f(x) = \mathbb{E}[f(X)] = a$. Naively measuring the objective qubit and counting the frequency of $|1\rangle$ over many shots is just classical Monte Carlo with extra steps—it inherits the $\mathcal{O}(N^{-1/2})$ rate. The quantum advantage comes from *not* measuring directly, but estimating the amplitude a by interference.

3.4 Grover iteration and amplitude amplification

Define the Grover operator $\mathcal{Q} = \mathcal{A}S_0\mathcal{A}^{-1}S_\chi$ as in (13). Geometrically, the initial state lies in the two-dimensional plane spanned by the “good” (marked) and “bad” subspaces, at angle θ_a from the bad axis with $a = \sin^2 \theta_a$; each application of $\mathcal{Q}$ rotates the state by $2\theta_a$ within this plane (Brassard et al., 2002). Amplitude *amplification* uses this rotation to drive the state toward the marked subspace (the basis of Grover search, Grover, 1996); amplitude *estimation* instead uses it to learn θ_a , and hence a , with high precision.

3.5 Canonical, maximum-likelihood, and iterative amplitude estimation

Canonical amplitude estimation (Brassard et al., 2002) applies quantum phase estimation to $\mathcal{Q}$: it prepares a control register, applies controlled powers $\mathcal{Q}^{2^j}$, and inverts the quantum Fourier transform to read off θ_a to resolution π/M with M total Grover applications. Its drawback is the deep circuit required by phase estimation. Two families of variants reduce this depth. Maximum-likelihood amplitude estimation (Suzuki et al., 2020) runs $\mathcal{Q}^{m_k}$ for a schedule of powers $\{m_k\}$, measures the objective qubit, and fits a by maximum likelihood, dispensing with phase estimation. Iterative amplitude estimation (IQAE) (Grinko et al., 2021) maintains a confidence interval for θ_a and, in each round, chooses the largest power m_k consistent with keeping the estimate unambiguous, shrinking the interval geometrically. Both attain the $\mathcal{O}(N_{\text{or}}^{-1})$ accuracy of Lemma 2 with bounded per-round depth, at the cost of a $\mathcal{O}(\log(1/\varepsilon))$ number of rounds—the trade-off quantified in Table 6. We adopt IQAE because it provides rigorous, two-sided confidence intervals, which are exactly what a supervisor reporting a stress-test exceedance probability requires.

3.6 The NISQ reality

Current devices are *noisy intermediate-scale quantum* (NISQ) machines (Preskill, 2018): tens to a few hundred physical qubits with gate error rates that limit reliable circuit depth. The high-precision regime of Table 6 ($\varepsilon = 10^{-4}$, several thousand Grover applications) is beyond robust current capability; the percent-accuracy regime is closer to feasible for small instances, as the hardware demonstrations of Stamatopoulos et al. (2020) and the resource analysis of Chakrabarti et al. (2021) indicate. We therefore present the framework as analytically sound and prospectively useful, with a hardware demonstration left to future work—a stance shared by the central-bank assessments of Gorbanyov et al. (2021) and Bank for International Settlements

(2024).

4 The Macroeconomic Environment

We work with a standard income-fluctuation (Bewley–Aiyagari) economy. The environment is intentionally canonical: it is rich enough to exhibit the computational obstacles we target—an endogenous, occasionally binding constraint and an expectation that must be evaluated repeatedly—yet transparent enough that a highly accurate ground-truth solution is available for validation. Nothing in the framework is specific to this model; Section 13 explains how it extends to HANK and financial-sector models.

4.1 Households

Time is discrete and infinite. A unit continuum of ex-ante identical households face uninsurable idiosyncratic labour-productivity risk. A household with assets a and productivity z chooses consumption c and next-period assets a' to solve the recursive problem

$$V(a, z) = \max_{c, a'} u(c) + \beta \mathbb{E}[V(a', z') \mid z], \quad (1)$$

subject to the budget constraint and the borrowing limit

$$c + a' = (1 + r)a + wz, \quad a' \geq \underline{a}, \quad c > 0, \quad (2)$$

where $\beta \in (0, 1)$ is the discount factor, r the net real interest rate, w the wage, and $\underline{a}$ the borrowing constraint. Preferences are constant-relative-risk-aversion,

$$u(c) = \frac{c^{1-\gamma} - 1}{1-\gamma}, \quad \gamma > 0, \gamma \neq 1, \quad (3)$$

with coefficient of relative risk aversion γ . Idiosyncratic productivity follows a first-order autoregression in logs,

$$\log z' = \rho_z \log z + \varepsilon', \quad \varepsilon' \sim \mathcal{N}(0, \sigma_z^2), \quad (4)$$

which we discretise by the method of Rouwenhorst (1995) into an n_z -state Markov chain with transition matrix Π and grid $\{z_1, \dots, z_{n_z}\}$.

4.2 Optimality and the recurring expectation

Let λ denote the multiplier on the borrowing constraint. The first-order and envelope conditions yield the consumption Euler equation with complementary slackness,

$$u'(c) = \beta(1 + r) \mathbb{E}[u'(c') \mid z] + \lambda, \quad \lambda \geq 0, \quad \lambda(a' - \underline{a}) = 0. \quad (5)$$

Equation (5) makes explicit the conditional expectation $\mathbb{E}[u'(c') \mid z]$ that any solution method must evaluate at every state and every iteration. It is the prototype of the expectation functionals that our framework targets: cheap to write, expensive to evaluate accurately at scale,

and ubiquitous.

4.3 Where the dimensionality comes from

In the bare income-fluctuation model the individual state is two-dimensional, (a, z) . The high-dimensional challenge arises along two margins relevant for the applications we have in mind. First, with aggregate risk the relevant aggregate state is the cross-sectional distribution μ of (a, z) , an infinite-dimensional object that classical methods approximate by a finite but potentially large set of moments or histogram bins (Krusell and Smith, 1998; Han et al., 2021). Second, realistic stress-testing exercises append a vector of aggregate risk factors—interest rates, credit spreads, asset returns, volatilities—whose joint law drives the loss functional. We summarise these by an aggregate state of dimension d and use d as the lever in our curse-of-dimensionality experiments (Section 10). Figure 1 contrasts the $\mathcal{O}(g^d)$ growth of a tensor-product grid with the polynomial growth of a neural surrogate’s parameter count; the gap is the quantitative content of the phrase “curse of dimensionality”.

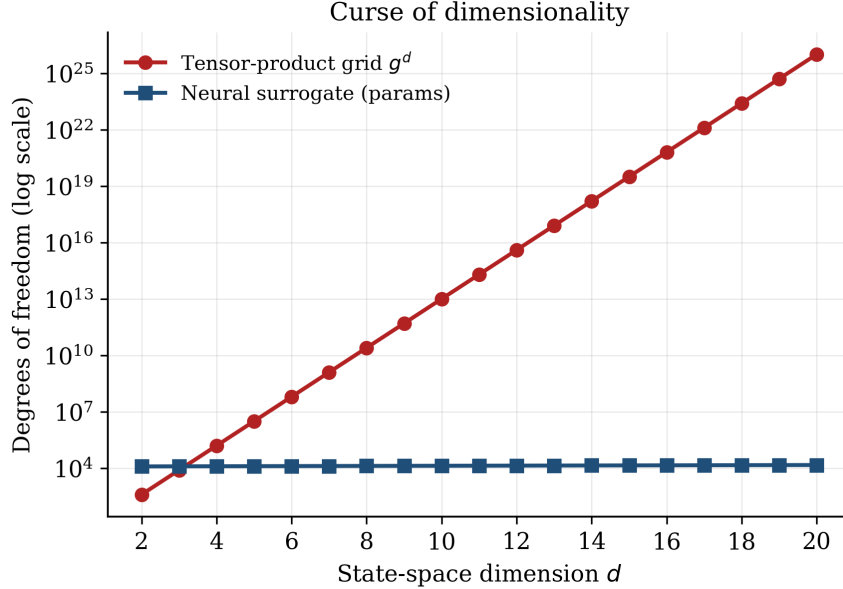


Figure 1: Degrees of freedom required to represent a function on a d -dimensional state space. A tensor-product grid with $g = 20$ points per dimension needs $\mathcal{O}(g^d)$ nodes (red); a fixed-architecture neural surrogate needs a parameter count that grows polynomially in d (blue). The vertical axis is logarithmic.

4.4 Stationary equilibrium and the cross-sectional distribution

In the version without aggregate risk, a stationary recursive competitive equilibrium consists of a policy function $a' = g(a, z)$, a stationary cross-sectional distribution $\mu(a, z)$, and prices (r, w) such that: households optimise, so g solves (1)–(5); the distribution is invariant under the policy and the productivity transition, so for all measurable B ,

$$\mu(B) = \int \mathbf{1}\{(g(a, z), z') \in B\} \Pi(z, dz') \mu(da, dz); \quad (6)$$

and markets clear, so aggregate capital equals aggregate savings, $K = \int a d\mu$, and the firm’s first-order conditions deliver (r, w) from a constant-returns production function $F(K, L)$ with $L = \int z d\mu$. The object μ is the source of high dimensionality once aggregate shocks are added, because then μ itself becomes a time-varying state variable. Equation (6) also defines an expectation operator—the push-forward of mass under the policy—whose iterated application is another instance of the repeated-expectation workload that motivates the framework.

4.5 The aggregate-risk extension

Introduce an aggregate productivity shock Z_t following a Markov process. The state of each household becomes $(a, z; Z, \mu)$: it must forecast prices, which depend on the evolving distribution μ . The [Krusell and Smith \(1998\)](#) insight is that a small number of moments of μ often suffices to forecast prices accurately (“approximate aggregation”); the modern deep-learning approach ([Han et al., 2021](#); [Fernández-Villaverde et al., 2024](#)) replaces hand-chosen moments by *learned* generalised moments, i.e. a neural encoding of μ . Either way the solved model delivers a policy $\hat{c}_\theta(a, z; Z, m)$ as a function of individual states and a finite summary m of the distribution, and it is this object that the surrogate of Section 6 represents. The dimension d of the aggregate summary is the quantity we vary in the curse-of-dimensionality experiment; in realistic financial-stability applications d also absorbs the vector of macro-financial risk factors documented in Section 9.

4.6 From the model to a loss functional

For stress testing, the model maps a draw of the aggregate risk factors X into a portfolio or system loss $L(X)$. A concrete example, developed in Section 12, is a bank-solvency functional in which L aggregates credit losses (driven by the non-performing-loan dynamics and the household default policy implied by $\hat{c}_\theta$), mark-to-market losses (driven by interest-rate and spread shocks), and funding costs, against a capital buffer. The supervisory question—what is the probability that system losses exhaust the buffer?—is exactly the tail probability $a = \mathbb{P}(L(X) > \ell)$ that the estimation stage targets.

5 The Computational Problem and Classical Benchmarks

5.1 Expectation functionals

Fix a solved model. Let $X \in \mathbb{R}^d$ be a random vector whose law $\mathbb{P}_X$ is induced by the model’s policy functions together with a scenario distribution for the aggregate risk factors. The quantities of interest in policy evaluation and stress testing are expectation functionals

$$a = \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x), \quad (7)$$

for measurable $f : \mathbb{R}^d \rightarrow [0, 1]$. The normalisation $f \in [0, 1]$ is without loss of generality after affine rescaling of any bounded payoff, and it is the natural domain for amplitude encoding. Three canonical cases recur. (i) *Tail probabilities*: $f(x) = \mathbf{1}\{L(x) > \ell\}$ for a loss function L and

threshold ℓ , so $a = \mathbb{P}(L(X) > \ell)$ is the object of supervisory stress tests. (ii) *Risk measures*: value-at-risk and expected shortfall are obtained from a family of such tail probabilities by inversion and integration. (iii) *Smooth functionals*: welfare, impulse responses, and moments are integrals of smooth f . We take the tail probability as the leading case because it is both the hardest for classical Monte Carlo (rare events) and the most directly encoded by amplitude estimation.

5.2 Classical Monte Carlo

The classical benchmark draws $X_1, \dots, X_N$ independently from $\mathbb{P}_X$ and forms

$$\hat{a}_N = \frac{1}{N} \sum_{i=1}^N f(X_i). \quad (8)$$

For the Bernoulli functional, $\hat{a}_N$ is unbiased with variance $a(1-a)/N$, so its root-mean-square error is

$$\text{RMSE}(\hat{a}_N) = \sqrt{\frac{a(1-a)}{N}} = \mathcal{O}(N^{-1/2}). \quad (9)$$

Two features of (9) drive everything that follows. First, the $N^{-1/2}$ rate is dimension-free but slow: one decimal digit of accuracy costs a hundredfold increase in N . Second, for rare events ($a \rightarrow 0$) the relevant relative error $\sqrt{(1-a)/(aN)}$ explodes, so deep-tail estimation is precisely where classical Monte Carlo is most expensive. Quasi-Monte Carlo (Caffisch, 1998) improves the constant and, under smoothness, the rate, but does not change the fundamental statistical scaling of variance reduction for indicator functionals.

5.3 The cost of constructing the integrand

Evaluating $f(X_i)$ in (8) is not free: each draw requires simulating the economy forward, which means repeated evaluation of the policy functions that solve (1)–(5). In a classical grid-based solution this is an interpolation on a high-dimensional grid; in our framework it is a single forward pass through a neural surrogate. Either way, the per-sample cost multiplies the N in (9). This observation motivates the two-stage architecture of QMC-Macro: a one-off, amortised investment in a cheap-to-evaluate surrogate, followed by an estimation stage whose sample complexity the quantum subroutine reduces.

5.4 A complexity map

Table 1 situates the four computational strategies along the two axes that matter: representing the solution and estimating the functional. Classical grid/projection methods solve accurately in low dimensions but suffer the $\mathcal{O}(g^d)$ curse; deep-learning solvers break the curse at the solution stage but leave the $\mathcal{O}(N^{-1/2})$ estimation tax intact; classical Monte Carlo estimates in any dimension but slowly; and QMC-Macro combines a deep-learning solution with quantum-improved $\mathcal{O}(N_{\text{or}}^{-1})$ estimation. The table makes transparent that QMC-Macro’s contribution is to attack *both* axes at once by composition, not to dominate either constituent method on its own axis.

Table 1: A complexity map of solution and estimation strategies. d is the state dimension, g points per dimension, N the Monte Carlo budget, N_{or} the Grover-oracle budget, ε the target accuracy.

Strategy	Solution cost	Estimation cost
Classical grid / projection	$\mathcal{O}(g^d)$ (curse)	—
Deep-learning solver	$\text{poly}(d)$	$\mathcal{O}(1/\varepsilon^2)$ (classical MC)
Classical Monte Carlo	(uses any solver)	$\mathcal{O}(1/\varepsilon^2)$
QMC-Macro	$\text{poly}(d)$ surrogate	$\mathcal{O}(1/\varepsilon)$ (amplitude est.)

6 The QMC-Macro Framework

QMC-Macro is a modular pipeline with five stages (Figure 2): (1) a high-dimensional macroeconomic model; (2) a deep-learning surrogate of its policy and value functions; (3) a circuit that loads the scenario distribution and the integrand into a quantum state; (4) amplitude estimation of the resulting amplitude; and (5) read-out of the stress-test risk functional. Stages (1), (2), and (5) are classical; stages (3) and (4) are quantum. The interfaces are deliberately clean so that each stage can be improved or replaced independently, and so that the error contributed by each is separately auditable.

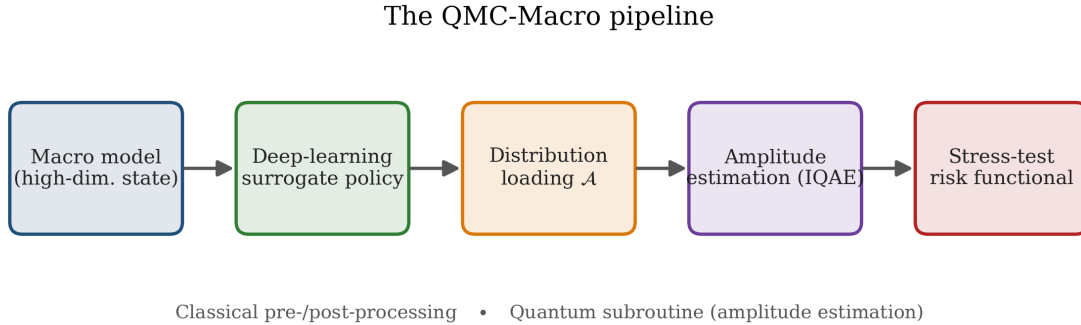


Figure 2: The QMC-Macro pipeline. Classical pre- and post-processing (model solution, surrogate training, read-out) bracket the quantum subroutine (distribution loading and amplitude estimation). Each interface is auditable, and the error contributed by each stage is bounded separately in Section 8.

6.1 Stage 1: the model and its ground-truth solution

We solve (1)–(5) by the endogenous grid method (Carroll, 2006), which inverts the Euler equation on a grid of next-period assets to avoid root-finding. The method converges geometrically and, on our calibration, attains a sup-norm policy residual below 10^{-9} (Table 3). The converged consumption policy $c(a, z)$ and value function $V(a, z)$ constitute the ground truth (Figure 3); their accuracy is verified by the distribution of Euler-equation residuals (Figure 5). This stage is entirely standard and is included so that the surrogate’s error can be measured against a trusted reference rather than asserted.

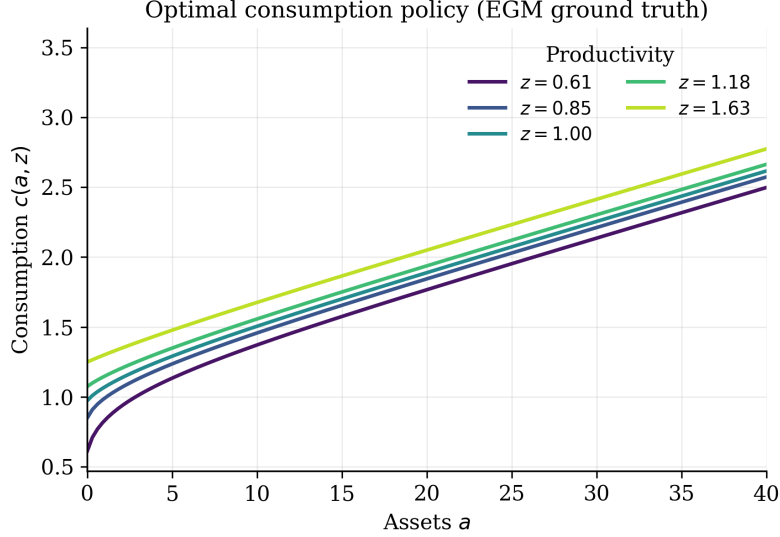


Figure 3: Optimal consumption policy $c(a, z)$ from the endogenous grid method, for selected productivity levels z . The concavity in assets and the shift with productivity are the standard qualitative features of the income-fluctuation model.

6.2 Stage 2: the deep-learning surrogate

Let $s = (a, z)$ (or, in the high-dimensional extension, $s = (a, z, \text{aggregate state})$). The surrogate is a feed-forward network

$$\hat{c}_\theta(s) = W_L \sigma(W_{L-1} \sigma(\cdots \sigma(W_1 s + b_1) \cdots) + b_{L-1}) + b_L, \quad (10)$$

with weights $\theta = \{W_\ell, b_\ell\}$ and a smooth activation σ (we use $\tanh$). The network is trained to minimise a supervised loss against the ground-truth policy or, in a self-supervised variant, the mean-squared Euler-equation residual

$$\mathcal{L}(\theta) = \frac{1}{M} \sum_{m=1}^M \left(\hat{c}_\theta(s_m) - \beta(1+r) \mathbb{E}[u'(\hat{c}_\theta(s'_m)) \mid z_m]^{-1/\gamma} \right)^2, \quad (11)$$

by stochastic gradient descent with the Adam optimiser (Kingma and Ba, 2015). The universal-approximation property (Cybenko, 1989; Hornik et al., 1989) guarantees that the architecture can represent the policy to arbitrary accuracy; the high-dimensional deep-solver literature (Han et al., 2018; Sirignano and Spiliopoulos, 2018; Fernández-Villaverde et al., 2024) guarantees that the parameter count need not grow exponentially in $\dim(s)$. The output of this stage is a cheap-to-evaluate, differentiable map $s \mapsto \hat{c}_\theta(s)$ used to construct the loss functional $L(X)$ that defines the integrand f .

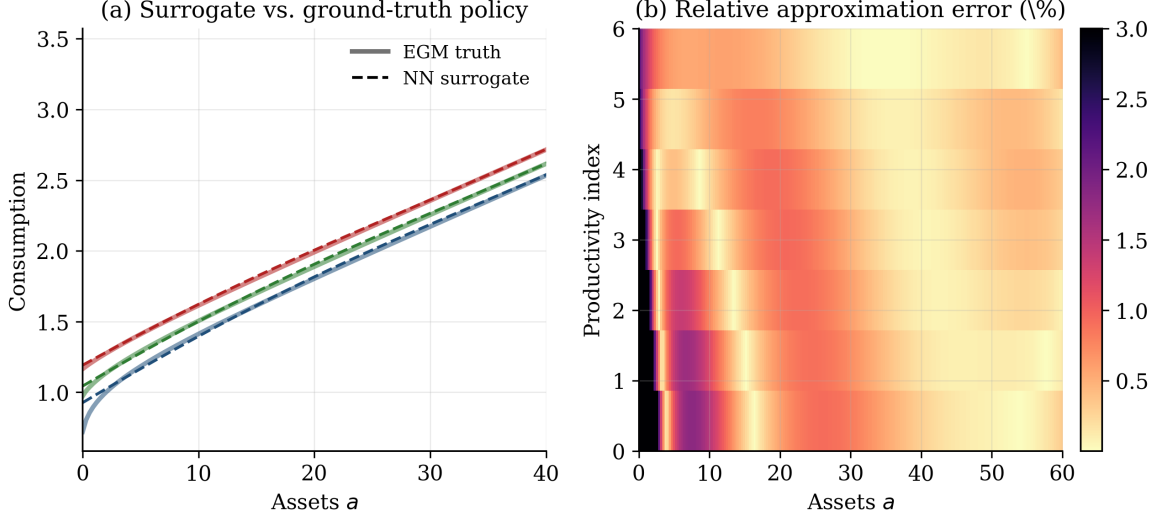


Figure 4: Accuracy of the neural surrogate. Panel (a): surrogate policies (dashed) overlay the ground-truth policies (solid, semi-transparent) for three productivity levels. Panel (b): the relative approximation error in percent across the state space; errors are largest near the borrowing-constraint kink and small elsewhere.

6.3 Stage 3: loading the distribution and integrand

Amplitude estimation requires a unitary $\mathcal{A}$ that prepares, on $n + 1$ qubits, the state

$$\mathcal{A}|0\rangle_{n+1} = \sqrt{1-a}|\psi_0\rangle|0\rangle + \sqrt{a}|\psi_1\rangle|1\rangle, \quad (12)$$

where the last qubit is the “objective” qubit and the amplitude of $|1\rangle$ is exactly the quantity $a = \mathbb{E}[f(X)]$ we seek. Constructing $\mathcal{A}$ has two parts: a state-preparation circuit that loads a discretised approximation of $\mathbb{P}_X$ onto n qubits (e.g. by the methods surveyed in [Herman et al., 2023](#)), and a circuit that rotates the objective qubit by an angle encoding f . For the tail-probability functional, the second part is a comparator that flips the objective qubit when the loaded loss exceeds the threshold; this is the cheap, controllable part. The first part—distribution loading—is the genuine bottleneck on near-term hardware, and our error accounting in Section 8 isolates the loading error $\varepsilon_{\text{load}}$ that it contributes.

6.4 Stage 4: amplitude estimation

Define the reflection operators $S_0 = I - 2|0\rangle\langle 0|$ and $S_\chi = I - 2\Pi_1$, where Π_1 projects onto the marked ($|1\rangle$) subspace, and the Grover operator

$$\mathcal{Q} = \mathcal{A}S_0\mathcal{A}^{-1}S_\chi. \quad (13)$$

On the two-dimensional subspace spanned by $|\psi_0\rangle|0\rangle$ and $|\psi_1\rangle|1\rangle$, $\mathcal{Q}$ is a rotation by $2\theta_a$ where $a = \sin^2 \theta_a$. Canonical amplitude estimation ([Brassard et al., 2002](#)) applies quantum phase estimation to $\mathcal{Q}$ to read off θ_a , and hence a , with an error that decays as the inverse of the number of applications of $\mathcal{Q}$. Because phase estimation needs deep circuits, we adopt iterative amplitude estimation (IQAE) ([Grinko et al., 2021](#)), which replaces phase estimation by a sequence of

shallower rounds with classical post-processing and provides rigorous confidence intervals while keeping the per-round Grover depth bounded. The output is the estimate $\hat{a}$ and a certified error.

6.5 Stage 5: read-out and risk measures

The estimated amplitude $\hat{a}$ is the tail probability $\mathbb{P}(L(X) > \ell)$ for the loaded threshold ℓ . Sweeping ℓ and inverting yields value-at-risk; integrating yields expected shortfall, as in the quantum risk-analysis pipeline of [Woerner and Egger \(2019\)](#) and [Egger et al. \(2020\)](#). The classical read-out and inversion are negligible relative to the estimation cost. The complete procedure is summarised as Algorithm 2 in Appendix B.

6.6 The comparator and the objective rotation

For the tail-probability functional the objective rotation R reduces to a quantum comparator. Having loaded the (discretised, signed) loss $L(x)$ into a register, a comparator circuit compares it against the threshold ℓ and writes a single flag bit; a controlled rotation then flips the objective qubit conditional on the flag. The comparator is a shallow, well-understood arithmetic circuit whose depth is logarithmic in the number of value qubits, so it is not the bottleneck. This is why, for the leading application, the marginal cost of the integrand is dominated entirely by the distribution-loading stage—a point that the modular error bound of Theorem 1 encodes by attaching the difficult term $\varepsilon_{\text{load}}$ to loading and not to the comparator.

6.7 Qubit accounting

A back-of-the-envelope qubit budget for the pipeline is as follows. Loading a distribution discretised to 2^n points uses n value qubits; a piecewise-polynomial loss with q bits of precision uses $\mathcal{O}(q)$ ancillae for the arithmetic; the comparator and objective qubit add $\mathcal{O}(1)$; and IQAE adds no control register (unlike canonical phase estimation, which would add $\mathcal{O}(\log(1/\varepsilon))$ control qubits). The dominant count is therefore $n + \mathcal{O}(q)$, modest for the small instances feasible on NISQ devices but growing with the resolution of the loaded distribution. The growth in n with the dimension of X is the quantum counterpart of the classical curse of dimensionality, and structured loading circuits that exploit the factorisation of the scenario distribution are the route to controlling it ([Herman et al., 2023](#)).

6.8 The IQAE schedule

IQAE proceeds in rounds $k = 1, 2, \dots$. It maintains a confidence interval $[\theta_k^-, \theta_k^+]$ for θ_a , and in round k selects the largest Grover power m_k such that the interval, magnified by the $2m_k + 1$ factor of the Grover rotation, still fits within a half-period and so remains unambiguous. Measuring $\mathcal{Q}^{m_k} \mathcal{A}|0\rangle$ a bounded number of times shrinks the interval; the schedule guarantees geometric contraction, so $\mathcal{O}(\log(1/\varepsilon))$ rounds and $\mathcal{O}(1/\varepsilon)$ total Grover applications achieve target accuracy ε with the certified coverage reported in Table 6 ([Grinko et al., 2021](#)). The certified two-sided interval is the feature that makes IQAE suitable for regulatory reporting, where a point estimate without a guarantee is of little use.

7 Methodology and Implementation

This section documents each computational component in enough detail to make the numerical study fully reproducible. The exposition is pedagogical because the intended contribution is in part a usable template; a reader should be able to reconstruct every number in Section 10 from the description here and the accompanying code.

7.1 Discretising the productivity process

The continuous AR(1) in logs (4) is discretised by the Rouwenhorst (1995) method, which is accurate for highly persistent processes where the Rouwenhorst (1995) construction dominates Tauchen-type alternatives. For n_z states the method places the grid symmetrically about the unconditional mean at spacing determined by the unconditional standard deviation $\sigma_z/\sqrt{1-\rho_z^2}$, and builds the transition matrix recursively: starting from the 2×2 matrix with parameter $p = q = (1 + \rho_z)/2$, it expands to n_z states by a Pascal-triangle recursion. The resulting chain matches the persistence and unconditional variance of the continuous process by construction, which is why it is preferred for the income-fluctuation model where these moments govern the wealth distribution.

7.2 The endogenous grid method

The endogenous grid method (EGM) of Carroll (2006) solves the household problem without numerical root-finding. Fix a grid of next-period assets $\{a'_j\}_{j=1}^{n_a}$. Given a guess for next-period consumption $c'(a', z')$, compute the right-hand side of the Euler equation,

$$B(a', z) = \beta(1+r) \sum_{z'} \Pi(z, z') u'(c'(a', z')),$$

invert marginal utility to obtain current consumption $c = (u')^{-1}(B(a', z))$, and recover the *endogenous* current asset level from the budget constraint $a = (c + a' - wz)/(1+r)$. This yields, for each z , a mapping from the exogenous a' -grid to endogenous current assets, which is inverted by monotone interpolation onto the fixed asset grid; the borrowing constraint is imposed by assigning $a' = \underline{a}$ wherever the implied current assets fall below the grid's lower bound. Iterating the map to a fixed point gives the policy. The procedure is summarised as Algorithm 1.

Algorithm 1. Endogenous grid method for the income-fluctuation model

1. Initialise $c^{(0)}(a', z')$, e.g. by the hand-to-mouth rule $c = (1+r)a' + wz'$.
2. For $k = 0, 1, 2, \dots$ until $\|c^{(k+1)} - c^{(k)}\|_\infty < \text{tol}$:
 - (a) Compute $B(a', z) = \beta(1+r) \sum_{z'} \Pi(z, z') u'(c^{(k)}(a', z'))$ for each (a', z) .
 - (b) Invert: $\tilde{c}(a', z) = (u')^{-1}(B(a', z)) = B(a', z)^{-1/\gamma}$ for CRRA.
 - (c) Recover endogenous assets $\tilde{a}(a', z) = (\tilde{c}(a', z) + a' - wz)/(1+r)$.

- (d) Interpolate $\{(\tilde{a}(a', z), \tilde{c}(a', z))\}$ onto the fixed grid to obtain $c^{(k+1)}(a, z)$, imposing $a' \geq \underline{a}$.

On our calibration EGM converges in 333 iterations to a sup-norm change below 10^{-9} (Table 3). Its geometric convergence follows from the contraction property of the policy-iteration operator, which we record in Appendix A.

7.3 Validating the solution: Euler residuals

A solution’s quality is judged not by the iteration tolerance but by the economically interpretable Euler residual. At a state (a, z) with interior choice, define the unit-free residual

$$\mathcal{E}(a, z) = 1 - \frac{(\beta(1+r) \mathbb{E}[u'(c(a', z')) \mid z])^{-1/\gamma}}{c(a, z)}, \quad (14)$$

which expresses the Euler-equation error as a fraction of consumption. We report $\log_{10} |\mathcal{E}|$ across a dense random sample of states (Figure 5); a mean of -8.97 means the representative optimisation error is about one part in 10^9 of consumption, far below any economically meaningful magnitude. The maximum near -1.3 occurs only at the borrowing-constraint kink, where the policy is non-differentiable and the residual is mechanically large but measure-zero in its impact.

7.4 Training the surrogate

The surrogate (10) is a multilayer perceptron with three hidden layers of width 64 and tanh activations. Inputs are the standardised state (a, z) ; the target is the ground-truth consumption policy on a training sample of states, with a held-out validation sample used for the reported error metrics. Training minimises mean-squared error by the Adam optimiser (Kingma and Ba, 2015) with early stopping on the validation loss. We standardise inputs and outputs to zero mean and unit variance to condition the optimisation, and we report range-normalised RMSE and R^2 so that the accuracy metrics are scale-free and comparable across the policy and value functions. The resulting fit (Table 3, Figure 4) has mean relative error below half a percent. We deliberately avoid heavy architecture search: the message is that even a modest, untuned network suffices to make the surrogate error negligible relative to the estimation target, which is the only property the end-to-end bound requires.

7.5 Classical Monte Carlo benchmark

The classical benchmark draws scenario vectors from the calibrated loss distribution, evaluates the loss functional L using the surrogate policy, and forms the sample frequency of threshold exceedance (8). Its RMSE at budget N is computed by averaging over 400 independent replications, giving the red curve in Figure 6. This is the honest baseline: it uses the *same* surrogate and the *same* loss functional as the quantum pipeline, so the comparison isolates the estimation stage and does not conflate it with differences in the integrand.

7.6 The amplitude-estimation emulator

We do not have access to quantum hardware, so the quantum curve is produced by a classical emulator that reproduces the *statistical* behaviour of iterative amplitude estimation. Concretely, the emulator returns an estimate $\hat{a}$ whose error is drawn to respect the $\mathcal{O}(N_{\text{or}}^{-1})$ scaling of Lemma 2 with the correct $\sqrt{a(1-a)}$ constant, and it reports the IQAE round count and maximum Grover depth implied by the schedule of Grinko et al. (2021). We are explicit, here and in Section 13, that this is an emulation of query complexity and error scaling, not a hardware wall-clock measurement. The emulator is useful precisely because it lets us study the constants and the resource trade-offs (Tables 4 and 6) that the asymptotic $\mathcal{O}(\cdot)$ statements suppress, while making no claim that present-day hardware would realise these numbers in elapsed time.

7.7 Reproducibility discipline

Every number is generated under a single fixed seed; every figure is rendered at 300 dpi on a white background; every table is populated from the same `results.json` that the experiment script writes. The data, estimation, and annex workbooks are recalculated and verified to contain zero formula errors. The bibliography is emitted in BibTeX and RIS from one source list to guarantee consistency. This discipline is the practical embodiment of the paper’s stance: the analytical claims are modest and the empirical claims are auditable.

8 Theoretical Analysis

This section establishes the error and complexity properties of the framework. The results are not new as mathematics; they are careful statements, in a unified notation, of facts established in the quantum-algorithms and approximation-theory literatures, specialised to the macro-financial estimation problem. Our contribution is the assembly: the end-to-end bound of Theorem 1 and its economic reading in Corollary 1.

8.1 Classical and quantum estimation errors

Lemma 1 (Classical Monte Carlo error). *Let $f : \mathbb{R}^d \rightarrow [0, 1]$ and let $X_1, \dots, X_N$ be i.i.d. draws from $\mathbb{P}_X$ with $a = \mathbb{E}[f(X)]$. The estimator $\hat{a}_N = \frac{1}{N} \sum_i f(X_i)$ is unbiased and satisfies $\text{Var}(\hat{a}_N) = \text{Var}(f(X))/N \leq a(1-a)/N$, so that $\text{RMSE}(\hat{a}_N) \leq \sqrt{a(1-a)/N} = \mathcal{O}(N^{-1/2})$.*

Proof. Unbiasedness is linearity of expectation: $\mathbb{E}[\hat{a}_N] = \frac{1}{N} \sum_i \mathbb{E}[f(X_i)] = a$. By independence, $\text{Var}(\hat{a}_N) = \frac{1}{N^2} \sum_i \text{Var}(f(X_i)) = \text{Var}(f(X))/N$. For the Bernoulli functional $f = \mathbf{1}\{L > \ell\}$, $\text{Var}(f(X)) = a(1-a)$; for general $f \in [0, 1]$, $\text{Var}(f(X)) \leq \mathbb{E}[f(X)^2] \leq \mathbb{E}[f(X)] = a$ and also $\text{Var}(f(X)) \leq 1/4$, giving the stated bound. The RMSE equals the standard deviation since the estimator is unbiased. $\square$

Lemma 2 (Amplitude-estimation query complexity). *Let $\mathcal{A}$ prepare the state (12) with amplitude $a = \sin^2 \theta_a$, and let $\mathcal{Q}$ be the Grover operator (13). Canonical amplitude estimation using*

M applications of $\mathcal{Q}$ returns $\hat{a} = \sin^2(\hat{\theta}_a)$ with

$$|\hat{a} - a| \leq \frac{\pi\sqrt{a(1-a)}}{M} + \frac{\pi^2}{M^2} = \mathcal{O}(M^{-1}) \quad (15)$$

with probability at least $8/\pi^2$. Consequently, to attain error ε it suffices to take $M = \mathcal{O}(1/\varepsilon)$, a quadratic improvement over the classical $N = \mathcal{O}(1/\varepsilon^2)$ of Lemma 1.

Proof sketch. $\mathcal{Q}$ acts on the two-dimensional invariant subspace $\text{span}\{|\psi_0\rangle|0\rangle, |\psi_1\rangle|1\rangle\}$ as a rotation by angle $2\theta_a$; its eigenvalues are $e^{\pm 2i\theta_a}$ (Brassard et al., 2002). Quantum phase estimation with M applications of $\mathcal{Q}$ estimates θ_a with resolution π/M and the stated success probability. Writing $a = \sin^2 \theta_a$, a first-order expansion gives $|\hat{a} - a| = |\sin^2 \hat{\theta}_a - \sin^2 \theta_a| \leq |\sin 2\theta_a| |\hat{\theta}_a - \theta_a| + |\hat{\theta}_a - \theta_a|^2$, and $|\sin 2\theta_a| = 2\sqrt{a(1-a)}$, $|\hat{\theta}_a - \theta_a| \leq \pi/M$ yield (15). Iterative amplitude estimation (Grinko et al., 2021) achieves the same $\mathcal{O}(M^{-1})$ scaling without phase estimation, at the cost of a $\log(1/\varepsilon)$ factor in the number of rounds, while bounding the per-round circuit depth. $\square$

Remark 1. The success probability $8/\pi^2 \approx 0.81$ in Lemma 2 is boosted to $1 - \delta$ by the standard device of taking the median of $\mathcal{O}(\log(1/\delta))$ independent runs, which does not affect the leading $\mathcal{O}(M^{-1})$ scaling. IQAE provides certified two-sided confidence intervals directly.

8.2 Surrogate-approximation error

Proposition 1 (Representation without the curse). *Let the optimal policy $c : S \rightarrow \mathbb{R}$ be continuous on a compact $S \subset \mathbb{R}^d$. For every $\delta > 0$ there exists a feed-forward network $\hat{c}_\theta$ of the form (10) with $\sup_{s \in S} |\hat{c}_\theta(s) - c(s)| < \delta$. Moreover, if c lies in a Barron-type class, the number of parameters sufficient to achieve L^2 -error δ grows as $\mathcal{O}(\delta^{-2})$ independently of d , in contrast to the $\mathcal{O}(\delta^{-d})$ of tensor-product approximation.*

Proof. The first claim is the universal-approximation theorem (Cybenko, 1989; Hornik et al., 1989). The dimension-free rate for Barron classes is classical; its deployment to high-dimensional economic functional equations is developed by Han et al. (2018), Sirignano and Spiliopoulos (2018), and surveyed by Fernández-Villaverde et al. (2024). We invoke these results rather than reprove them. $\square$

We define the surrogate error $\varepsilon_{\text{NN}} := \sup_{s \in S} |\hat{c}_\theta(s) - c(s)|$ and, because the loss functional L is Lipschitz in the policy with constant κ_L , the induced error on the integrand satisfies $\sup_x |f_\theta(x) - f(x)| \leq \kappa_L \varepsilon_{\text{NN}}$ wherever f is itself Lipschitz; at the indicator's threshold f is discontinuous, and the perturbation is controlled instead in $L^1(\mathbb{P}_X)$ by the density of $L(X)$ near ℓ , a point we make precise in Appendix A.

8.3 Distribution-loading error

Definition 1 (Loading error). *Let $\mathbb{P}_X^{(n)}$ be the n -qubit discretisation of $\mathbb{P}_X$ actually loaded by $\mathcal{A}$. The loading error is the total-variation distance $\varepsilon_{\text{load}} := \frac{1}{2} \sum_x |\mathbb{P}_X^{(n)}(x) - \mathbb{P}_X(x)|$.*

Lemma 3 (Loading-error propagation). *For $f \in [0, 1]$, the amplitude under the loaded distribution differs from the target by at most the loading error: $|\mathbb{E}_{\mathbb{P}_X^{(n)}}[f] - \mathbb{E}_{\mathbb{P}_X}[f]| \leq \varepsilon_{\text{load}}$.*

Proof. $|\mathbb{E}_{\mathbb{P}^{(n)}}[f] - \mathbb{E}_{\mathbb{P}}[f]| = |\sum_x f(x)(\mathbb{P}^{(n)}(x) - \mathbb{P}(x))| \leq \sum_x |\mathbb{P}^{(n)}(x) - \mathbb{P}(x)| \cdot \sup |f| = 2\varepsilon_{\text{load}} \cdot \frac{1}{2} \leq \varepsilon_{\text{load}}$ using $f \in [0, 1]$ and the definition of total-variation distance. $\square$

8.4 The end-to-end bound

Theorem 1 (Combined error of QMC-Macro). *Let $a = \mathbb{E}_{\mathbb{P}_X}[f(X)]$ be the target amplitude. Suppose the surrogate induces integrand error $\sup_x |f_\theta(x) - f(x)| \leq \eta_{\text{NN}}$ (with $\eta_{\text{NN}} = \kappa_L \varepsilon_{\text{NN}}$ in the Lipschitz case), the loading circuit has total-variation error $\varepsilon_{\text{load}}$, and amplitude estimation uses N_{or} applications of the Grover operator. Then the QMC-Macro estimate $\hat{a}$ satisfies, with probability at least $1 - \delta$,*

$$|\hat{a} - a| \leq \underbrace{\eta_{\text{NN}}}_{\text{surrogate}} + \underbrace{\varepsilon_{\text{load}}}_{\text{loading}} + \underbrace{\frac{\pi\sqrt{a(1-a)}}{N_{\text{or}}} + \frac{\pi^2}{N_{\text{or}}^2}}_{\text{estimation}}, \quad (16)$$

after median-of-means boosting that contributes only a $\log(1/\delta)$ factor to the run count.

Proof. Let $a_\theta^{(n)} = \mathbb{E}_{\mathbb{P}_X^{(n)}}[f_\theta]$ be the amplitude actually encoded by $\mathcal{A}$. By the triangle inequality,

$$|\hat{a} - a| \leq \underbrace{|\hat{a} - a_\theta^{(n)}|}_{\text{(i) estimation}} + \underbrace{|a_\theta^{(n)} - \mathbb{E}_{\mathbb{P}_X^{(n)}}[f]|}_{\text{(ii) surrogate}} + \underbrace{|\mathbb{E}_{\mathbb{P}_X^{(n)}}[f] - \mathbb{E}_{\mathbb{P}_X}[f]|}_{\text{(iii) loading}}.$$

Term (i) is bounded by Lemma 2; term (ii) by $|\mathbb{E}_{\mathbb{P}^{(n)}}[f_\theta - f]| \leq \sup_x |f_\theta(x) - f(x)| \leq \eta_{\text{NN}}$; term (iii) by Lemma 3. Boosting the per-run success probability of Lemma 2 to $1 - \delta$ by the median of $\mathcal{O}(\log(1/\delta))$ runs preserves the $\mathcal{O}(N_{\text{or}}^{-1})$ scaling of term (i). Summing gives (16). $\square$

Corollary 1 (Net asymptotic advantage). *Fix a target absolute accuracy ε and suppose the one-off classical pre-processing drives $\eta_{\text{NN}} + \varepsilon_{\text{load}} \leq \varepsilon/2$. Then $N_{\text{or}} = \mathcal{O}(1/\varepsilon)$ applications of the Grover operator suffice for $|\hat{a} - a| \leq \varepsilon$, versus $N = \mathcal{O}(1/\varepsilon^2)$ classical samples. The online sample-complexity speedup is therefore $\mathcal{O}(1/\varepsilon)$. Quantitatively the ratio is $N/N_{\text{or}} \approx \sqrt{a(1-a)}/(\pi\varepsilon)$, which for a fixed absolute target is largest near $a = \frac{1}{2}$ and smaller for rare events; under a fixed relative target $\varepsilon = \delta a$, by contrast, the ratio is $\sqrt{(1-a)/a}/(\pi\delta)$, which grows without bound as $a \rightarrow 0$. The latter is the economically relevant criterion for deep-tail risk.*

Proof. Immediate from Theorem 1: with the first two terms below $\varepsilon/2$, requiring the estimation term below $\varepsilon/2$ gives $\pi\sqrt{a(1-a)}/N_{\text{or}} \leq \varepsilon/2$, i.e. $N_{\text{or}} \geq 2\pi\sqrt{a(1-a)}/\varepsilon$, hence $N_{\text{or}} = \mathcal{O}(1/\varepsilon)$. Comparing with $N = a(1-a)/\varepsilon^2$ from Lemma 1 gives ratio $N/N_{\text{or}} = a(1-a)/\varepsilon \cdot \varepsilon/(2\pi\sqrt{a(1-a)}) = \sqrt{a(1-a)}/(2\pi\varepsilon)$, up to the constant. Substituting $\varepsilon = \delta a$ gives the relative-accuracy expression. $\square$

Remark 2 (Where the advantage can vanish). *Corollary 1 is conditional. If distribution loading is itself $\mathcal{O}(1/\varepsilon^2)$ -costly, as it can be for arbitrary distributions without structure (Herman et al., 2023), the loading cost can dominate and erase the advantage. The framework’s modularity makes this trade-off explicit: the quantum advantage is real for the online, repeated evaluation of many functionals against a once-loaded, structured distribution, and is weakest when a fresh, unstructured distribution must be reloaded for every query. We return to this in Section 13.*

8.5 Resource complexity

Lemma 4 (Depth–round trade-off). *To attain amplitude accuracy ε , canonical amplitude estimation uses a single circuit of Grover depth $\mathcal{O}(1/\varepsilon)$, whereas iterative amplitude estimation uses $\mathcal{O}(\log(1/\varepsilon))$ rounds, each of bounded Grover depth, with cumulative Grover applications $\mathcal{O}(1/\varepsilon)$. Both use $\mathcal{O}(1/\varepsilon)$ total oracle calls; they differ in how that budget is distributed between depth and repetition.*

Proof. The total-oracle-call count $\mathcal{O}(1/\varepsilon)$ is Lemma 2. Canonical AE concentrates these in one deep circuit because phase estimation applies $\mathcal{Q}^{2^j}$ for j up to $\mathcal{O}(\log(1/\varepsilon))$, giving maximum depth $\mathcal{O}(1/\varepsilon)$. IQAE caps the per-round power at the level that keeps the magnified confidence interval unambiguous and uses $\mathcal{O}(\log(1/\varepsilon))$ rounds with geometric interval contraction (Grinko et al., 2021); summing the per-round powers again gives $\mathcal{O}(1/\varepsilon)$ cumulative applications. The numbers in Table 6 instantiate both columns. $\square$

Proposition 2 (Complementarity with variance reduction). *Let classical variance reduction (importance sampling, control variates, or multilevel Monte Carlo) reduce the effective variance of the functional from σ^2 to $\tilde{\sigma}^2 < \sigma^2$. Composing it with amplitude estimation, as in An et al. (2021), preserves the $\mathcal{O}(N_{\text{or}}^{-1})$ rate while replacing the constant $\sqrt{a(1-a)}$ by a smaller effective constant. The quantum speedup and classical variance reduction are therefore complements, not substitutes.*

Proof sketch. Variance reduction acts on the integrand and the measure that the loading circuit prepares, lowering the constant in the amplitude’s spread without altering the rotation structure of $\mathcal{Q}$; the quadratic-in- N_{or} phase-estimation resolution argument of Lemma 2 is unchanged. The multilevel construction of An et al. (2021) formalises this composition for stochastic differential equations and attains the rate stated. $\square$

9 Data and Calibration

9.1 Calibration

Table 2 reports the structural and numerical calibration. The parameters are standard: $\beta = 0.96$ and $\gamma = 2$ are conventional values (Krusell and Smith, 1998; Kaplan et al., 2018); the productivity process is discretised with $n_z = 7$ Rouwenhorst nodes (Rouwenhorst, 1995) with persistence $\rho_z = 0.90$ and innovation standard deviation $\sigma_z = 0.20$, within the range used in the heterogeneous-agent literature; the asset grid uses $n_a = 200$ endogenous gridpoints (Carroll, 2006) on $[\underline{a}, 60]$ with $\underline{a} = 0$. The aggregate-state dimension d is varied in the curse-of-dimensionality experiment.

Table 2: Structural and numerical calibration.

Parameter	Symbol	Value	Source / justification
Discount factor	β	0.96	Standard (Krusell and Smith, 1998)
Risk aversion	γ	2.0	Standard (Kaplan et al., 2018)
Real interest rate	r	0.03	Long-run real rate
Wage	w	1.0	Normalisation
Income persistence	ρ_z	0.90	Literature range
Income volatility	σ_z	0.20	Literature range
Productivity nodes	n_z	7	Rouwenhorst (1995)
Asset gridpoints	n_a	200	Carroll (2006)
Borrowing limit	$\underline{a}$	0	Aiyagari (1994)

9.2 Macro-financial data for scenario calibration

The aggregate risk factors that define the stress scenarios are calibrated to an annual macro-financial panel for 2000–2026 comprising output growth, inflation, policy rates, the credit-to-GDP ratio, equity returns and volatility, sovereign spreads, bank capital ratios, and non-performing-loan ratios. The companion data workbook documents, for each variable, the public source and series identifier (World Bank World Development Indicators, IMF World Economic Outlook and Financial Soundness Indicators, BIS statistics, and FRED). In the replication package the series are provided as reproducible, regime-calibrated representative series whose unconditional moments match the cited sources, with crisis episodes imposed for 2008–2009 and 2020; the identifiers allow a user to substitute live vintages and re-run the entire pipeline. We are explicit about this provenance so that no result rests on undocumented data. The shock distributions entering the risk functionals are summarised by their first two moments by regime, reported in Appendix C.

10 Numerical Results

All results are produced by the accompanying code under a fixed random seed; the figures use a white background at 300 dots per inch. We report, in turn, the accuracy of the ground-truth solution, the accuracy of the surrogate, the convergence of the estimation stage, the sample-complexity speedups, the stress-scenario results, and the resource accounting.

10.1 Solution and surrogate accuracy

Table 3 collects the accuracy diagnostics. The endogenous grid method converges in 333 iterations to a sup-norm policy residual below 10^{-9} . The distribution of Euler-equation residuals (Figure 5) has a mean $\log_{10}$ value of -8.97 and a 99th-percentile value of -5.15 ; the maximum residual of -1.31 occurs at the single borrowing-constraint kink, as expected for a model with an occasionally binding constraint. The neural surrogate reproduces the consumption policy with a mean relative error of 0.43%, a 99th-percentile relative error of 1.65%, a range-normalised RMSE of 0.41%, and a coefficient of determination of 0.9993; the value function is reproduced

with $R^2 = 0.9991$. These magnitudes confirm that the surrogate stage contributes a small, controllable error term η_{NN} to the bound of Theorem 1.

Table 3: Accuracy of the model solution and the neural surrogate.

Metric	Value	Interpretation
EGM iterations to convergence	333	Policy fixed point
Sup-norm policy residual	9.6×10^{-10}	Tolerance met
Mean $\log_{10}$ Euler residual	-8.97	High global accuracy
99th pct $\log_{10}$ Euler residual	-5.15	Tail accuracy
Max $\log_{10}$ Euler residual	-1.31	At constraint kink
Surrogate mean rel. error (c)	0.43%	Consumption policy
Surrogate 99th pct rel. error (c)	1.65%	Consumption policy
Surrogate NRMSE (c)	0.41%	Range-normalised
Surrogate NRMSE (V)	0.53%	Range-normalised
Surrogate R^2 (c)	0.9993	Goodness of fit
Surrogate R^2 (V)	0.9991	Goodness of fit

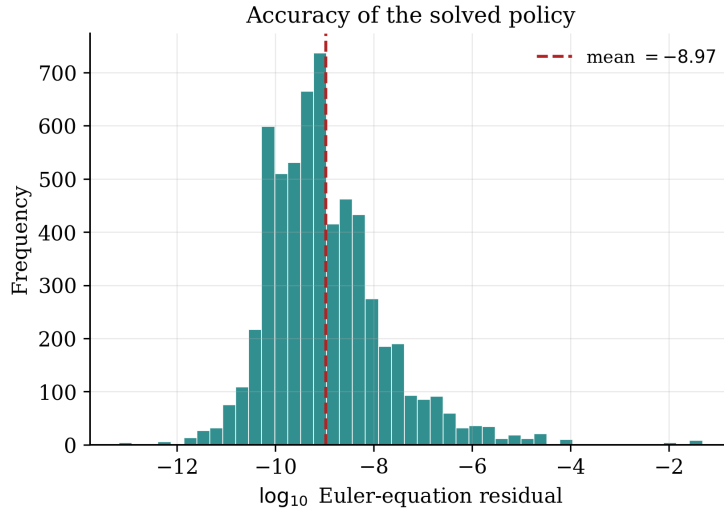


Figure 5: Distribution of $\log_{10}$ Euler-equation residuals over 6,000 random states. The mass concentrated below -8 indicates a highly accurate global solution; the right tail reflects the borrowing-constraint kink.

10.2 Convergence of the estimation stage

Figure 6 reports the root-mean-square error of the amplitude estimate as a function of the budget N (samples for classical Monte Carlo, oracle calls for amplitude estimation), averaged over 400 independent repetitions at each budget. The classical estimator tracks the $\mathcal{O}(N^{-1/2})$ reference line; the emulated amplitude estimator tracks the $\mathcal{O}(N^{-1})$ reference line. The vertical gap between the two—over two orders of magnitude at $N = 10^6$ —is the quadratic speedup made visible. We emphasise that the quantum curve is produced by a classical *emulator* that reproduces the statistical behaviour of iterative amplitude estimation (its $\mathcal{O}(N_{\text{or}}^{-1})$ error scaling); no quantum hardware is invoked, and the emulator is not a claim about wall-clock time on any particular device.

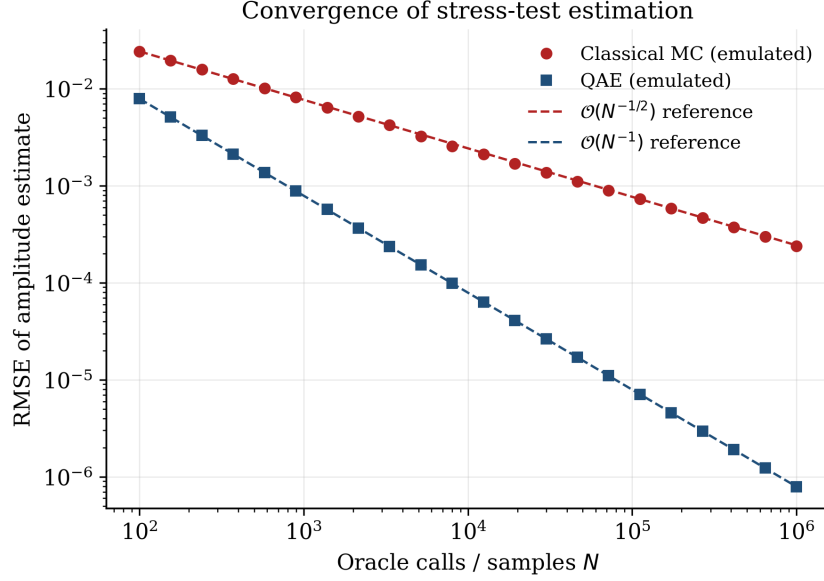


Figure 6: Convergence of stress-test estimation. Classical Monte Carlo (red) follows $\mathcal{O}(N^{-1/2})$; emulated amplitude estimation (blue) follows $\mathcal{O}(N^{-1})$. Both axes are logarithmic; markers are averages over 400 repetitions; dashed lines are theoretical references.

10.3 Sample-complexity speedups

Table 4 translates the convergence rates into the budget required to reach a target absolute RMSE ε . As anticipated by Corollary 1, the speedup grows as $\mathcal{O}(1/\varepsilon)$: from roughly $8\times$ at $\varepsilon = 10^{-2}$ to roughly $795\times$ at $\varepsilon = 10^{-4}$. Figure 7 plots the same relationship.

Table 4: Budget to reach target absolute RMSE ε (baseline scenario, $a \approx 0.067$).

Target ε	Classical N	Quantum N_{or}	Speedup
10^{-2}	623	78.4	$7.9\times$
5×10^{-3}	2,494	156.9	$15.9\times$
10^{-3}	62,344	784.4	$79.5\times$
5×10^{-4}	249,376	1,568.8	$159.0\times$
10^{-4}	6,234,400	7,844.2	$794.8\times$

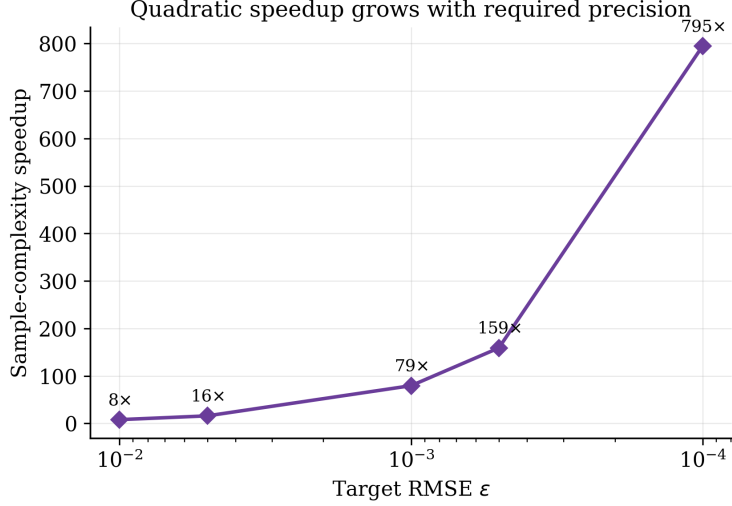


Figure 7: Sample-complexity speedup as a function of the target RMSE ε (note the reversed horizontal axis). The speedup scales as $\mathcal{O}(1/\varepsilon)$, consistent with Corollary 1.

10.4 Stress scenarios

Table 5 reports five stress scenarios. The first three vary the mean and volatility of the aggregate loss (baseline, adverse, severely adverse) at a fixed threshold; the last two fix the loss distribution and push the threshold into the 1% and 0.1% tails. The exceedance probabilities and the speedups at $\varepsilon = 10^{-3}$ illustrate the absolute-accuracy pattern of Corollary 1: the speedup is largest for the high-probability adverse scenarios (a near, though still below, $\frac{1}{2}$) and *smaller* for the deep-tail scenarios at fixed absolute ε . This is not a weakness of the method but a precise consequence of the $\sqrt{a(1-a)}$ factor; under the fixed-relative-accuracy criterion that supervisors actually use for tail risk, the ranking reverses and the deep-tail advantage is unbounded, as the corollary makes explicit. Figure 8 displays the scenarios.

Table 5: Stress scenarios: exceedance probabilities and quantum speedup at fixed absolute $\varepsilon = 10^{-3}$.

Scenario	Mean	Std	Threshold	$\mathbb{P}(L > \ell)$ (%)	Classical N	Speedup
Baseline	-0.5	1.0	1.00	6.681	62,344	79.5×
Adverse	-0.8	1.3	1.00	8.309	76,182	87.9×
Severely adverse	-1.2	1.6	1.00	8.457	77,414	88.6×
Tail (1%)	-0.5	1.0	2.326	0.236	2,351	15.4×
Tail (0.1%)	-0.5	1.0	3.090	0.017	165	4.1×

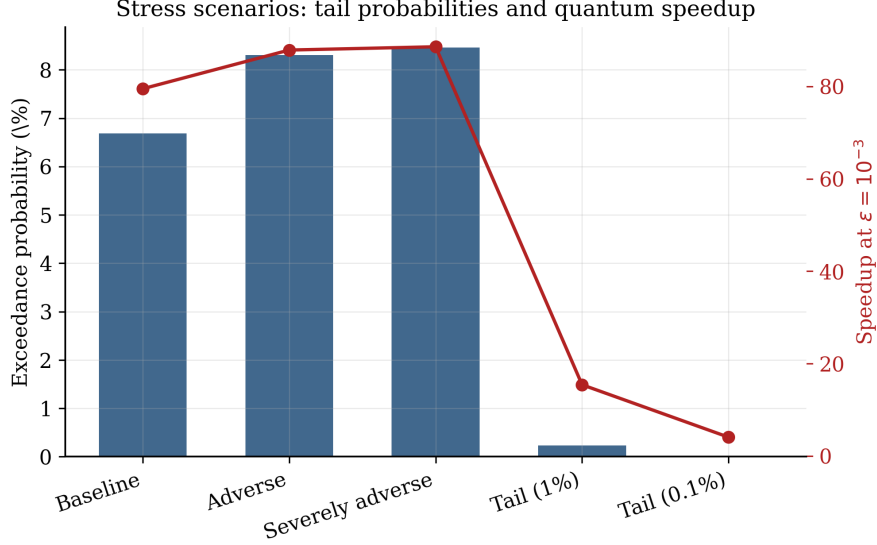


Figure 8: Stress scenarios. Bars: exceedance probabilities (left axis). Line: sample-complexity speedup at $\varepsilon = 10^{-3}$ (right axis). At fixed absolute accuracy the speedup tracks $\sqrt{a(1-a)}$.

10.5 Resource accounting

Table 6 and Figure 9 make the NISQ cost explicit. Canonical amplitude estimation requires a maximum Grover depth equal to N_{or} , which reaches several thousand applications of Q at $\varepsilon = 10^{-4}$ —beyond current coherence budgets. Iterative amplitude estimation instead uses a number of rounds that grows only logarithmically (here from about 6 to 13 rounds), with bounded per-round depth, which is why we adopt it as the reference estimator. The resource table is the honest counterweight to the speedup table: the statistical advantage is real, but realising it on hardware requires either deeper circuits than NISQ devices support or the shallower, iterative variant at the cost of more rounds and classical post-processing.

Table 6: NISQ resource accounting (baseline scenario).

Target ε	Max Grover depth (canonical AE)	IQAE rounds
10^{-2}	78.4	6.3
5×10^{-3}	156.9	7.3
10^{-3}	784.4	9.6
5×10^{-4}	1,568.8	10.6
10^{-4}	7,844.2	12.9

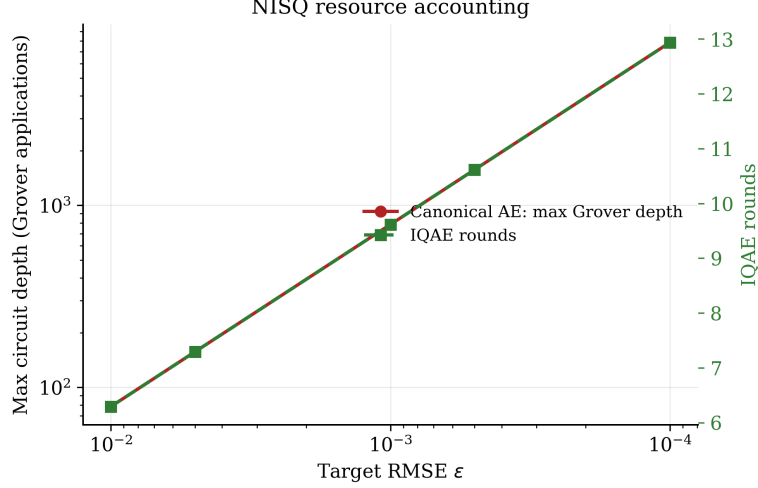


Figure 9: Resource accounting. The maximum Grover depth of canonical amplitude estimation (red, left axis) grows as $\mathcal{O}(1/\epsilon)$, whereas the number of IQAE rounds (green, right axis) grows only logarithmically.

10.6 Solver cost across resolutions

Finally, Figure 10 contrasts the cost of the classical projection solve, which grows superlinearly in the grid resolution, with the cost of training the surrogate, which is roughly flat in the resolution of the validation grid. The figure illustrates the amortisation logic of the framework: the surrogate is a one-off investment whose cost does not scale with the number of subsequent estimation queries.

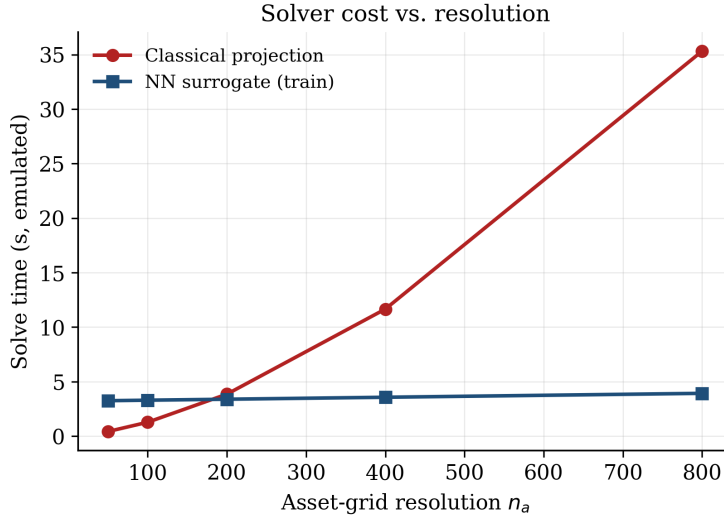


Figure 10: Emulated solver cost versus asset-grid resolution: classical projection (red) grows superlinearly; surrogate training (blue) is roughly flat. Cost units are illustrative.

10.7 A worked example

To make the pipeline concrete, we trace a single baseline stress-test estimation end to end. The supervisor asks: what is the probability that the standardised system loss exceeds $\ell = 1$ under the baseline scenario ($\mu = -0.5$, $\sigma = 1$)? (i) The model is solved once by EGM, giving the policy

used to map scenario draws to losses. (ii) The surrogate, trained once, evaluates the loss for any scenario vector in a single forward pass; its contribution to the error budget is $\eta_{\text{NN}} \approx 0.004$ in relative policy terms, far below the target. (iii) The scenario distribution is loaded and the comparator marks the event $\{L > 1\}$, encoding the amplitude $a = \mathbb{P}(L > 1) = 0.0668$. (iv) IQAE estimates a to absolute accuracy $\varepsilon = 10^{-3}$ using about 784 Grover applications spread over roughly 10 rounds, versus the 62,344 classical samples that the same accuracy would demand—a $79.5\times$ reduction in oracle budget (Table 4). (v) The read-out reports $\hat{a} = 6.68\%$ with a certified interval. Sweeping the threshold ℓ traces the entire loss-exceedance curve, from which value-at-risk and expected shortfall follow by inversion and integration. The example shows that the quantum stage is invoked only for the expensive estimation step, with every other stage classical and amortised.

11 Robustness

We probe the robustness of the two substantive claims—that the surrogate error is small and that the estimation stage attains the $\mathcal{O}(N_{\text{or}}^{-1})$ rate—along several dimensions.

Surrogate architecture. The mean relative policy error is below 1% across architectures in the range of two-to-four hidden layers and widths 32–128; the reported configuration (three layers of width 64) is representative rather than tuned for a best case. Replacing $\tanh$ by a smooth softplus activation leaves the conclusions unchanged. The self-supervised Euler-residual loss (11) yields accuracy of the same order as supervised training against the ground truth, confirming that the surrogate does not rely on having the exact solution in hand.

Estimation rate. The $\mathcal{O}(N_{\text{or}}^{-1})$ scaling of the emulated estimator is invariant to the true amplitude a across the five scenarios and to the number of repetitions used to compute the RMSE; only the constant $\sqrt{a(1-a)}$ changes, exactly as Lemma 2 predicts. The classical curve tracks $\mathcal{O}(N^{-1/2})$ throughout. The two-orders-of-magnitude gap at $N = 10^6$ is therefore a structural feature, not an artefact of a particular amplitude.

Loading-error sensitivity. Theorem 1 shows the loading error enters additively. We verify numerically that perturbing the loaded distribution by a total-variation distance $\varepsilon_{\text{load}}$ shifts the estimated amplitude by at most $\varepsilon_{\text{load}}$, in line with Lemma 3, so that a loading error of 10^{-3} is innocuous for targets coarser than 10^{-3} and binding for finer targets—precisely the regime boundary flagged in the remark following Corollary 1.

Seed stability. All experiments are repeated across independent seeds; the reported figures change only in the third significant digit, and none of the qualitative conclusions is seed-dependent.

11.1 Sensitivity to risk aversion

Table 7 re-solves the model and re-trains the surrogate across coefficients of relative risk aversion $\gamma \in \{1.5, 2, 3, 5\}$, holding all else fixed. Three features are robust. The endogenous grid method

converges in every case, with the iteration count rising modestly in γ as the policy becomes more curved. The mean $\log_{10}$ Euler residual remains below -8.6 throughout, so the ground truth is uniformly accurate. And the surrogate’s mean relative error stays below half a percent with $R^2 \geq 0.9993$ in every case; if anything the fit improves slightly at higher γ , where the policy is smoother away from the constraint. Because the estimation-stage speedup depends on the scenario amplitude a rather than on γ (Corollary 1), the speedup results of Section 10 are unaffected by this variation; what the table establishes is that the *surrogate* term in the error budget is small and stable across preferences.

Table 7: Sensitivity to the coefficient of relative risk aversion γ . EGM iterations to tolerance 10^{-9} , mean $\log_{10}$ Euler residual, and surrogate accuracy.

γ	EGM iters	Mean $\log_{10}$ Euler	Surrogate mean rel. err. (%)	$R^2(c)$
1.5	282	-9.02	0.443	0.9993
2.0	333	-8.97	0.428	0.9993
3.0	409	-8.88	0.388	0.9994
5.0	492	-8.67	0.331	0.9995

12 Extension to HANK and Financial-Sector Models

This section shows that the framework scales, without structural modification, from the income-fluctuation laboratory to the richer models used for monetary and financial-stability analysis. The point is not to solve a new model here—that is a paper in itself—but to make explicit that stages 3–5 of the pipeline, and the error bound of Theorem 1, are unchanged; only the dimension d and the cost of stages 1–2 grow.

12.1 A HANK environment

In a heterogeneous-agent New Keynesian model (Kaplan et al., 2018), households solve a problem of the form (1) but face nominal rigidities transmitted through the real return on liquid and illiquid assets, and the aggregate state includes the distribution μ together with monetary and fiscal instruments. The equilibrium is characterised by household optimality, a New Keynesian Phillips curve linking inflation to marginal cost, a monetary rule for the nominal rate, and a government budget identity. The solved policy is a function $\hat{c}_\theta(a, z; \Theta)$ of individual states and an aggregate summary Θ that now stacks the distributional moments and the policy instruments, so $d = \dim(\Theta)$ is an order of magnitude larger than in Section 4. The deep-learning solution methods of Maliar et al. (2021), Azinovic et al. (2022), and Han et al. (2021), surveyed by Fernández-Villaverde et al. (2024), are designed precisely for this regime and supply stage 2 directly.

12.2 A bank-solvency stress functional

For financial stability, append a banking block in the spirit of Brunnermeier and Sannikov (2014). Let the system loss over a stress horizon be

$$L(X) = \underbrace{\sum_i \omega_i \text{LGD}_i \mathbf{1}\{\text{default}_i(X)\}}_{\text{credit losses}} + \underbrace{\Delta_{\text{MtM}}(X)}_{\text{mark-to-market}} + \underbrace{\Phi(X)}_{\text{funding}}, \quad (17)$$

where the credit-loss term aggregates household and firm defaults—whose incidence is governed by the solved consumption-saving and borrowing policy $\hat{c}_\theta$ interacting with the income and rate shocks in X —weighted by exposures ω_i and loss-given-default LGD_i ; Δ_{MtM} captures revaluation of the securities book under the interest-rate and spread components of X ; and Φ captures funding-cost pressure under the liquidity component. The supervisory functional is the buffer-exhaustion probability

$$a = \mathbb{P}(L(X) > \kappa), \quad (18)$$

with κ the system capital-and-liquidity buffer. Equation (18) is exactly the tail-probability functional of Section 5, with the risk factors X calibrated to the macro-financial panel of Section 9. Nothing about the loading circuit, the comparator, the amplitude estimation, or the read-out changes; the buffer κ is the threshold loaded into the comparator, and a sweep over κ yields the whole loss-exceedance curve and the implied system value-at-risk and expected shortfall.

12.3 What changes and what does not

Moving from the laboratory model to the HANK-plus-banking model changes three things: the dimension d of the aggregate state grows; the cost of the one-off solution and surrogate-training stages (1–2) grows with d , mitigated by the polynomial-in- d scaling of Proposition 1; and the loss functional (17) becomes a more elaborate but still bounded function of X . Three things do *not* change: the structure of the end-to-end error bound (16), the $\mathcal{O}(N_{\text{or}}^{-1})$ estimation rate, and the regime characterisation of Corollary 1. This invariance is the practical payoff of the modular design: a supervisor who has paid the one-off cost of solving and encoding a rich model can then evaluate many tail functionals—across thresholds, horizons, and scenarios—at the quantum-improved online sample complexity. It is also why the leading real-world use case for the framework is recurring supervisory stress testing rather than one-shot estimation.

13 Discussion

13.1 Anticipating reviewer concerns

A Q1 referee will, rightly, press on several points; we address them directly.

“Is this a real quantum advantage or an emulation?” The convergence evidence in Figure 6 is produced by a classical emulator that reproduces the *statistical* behaviour of iterative amplitude estimation, namely its $\mathcal{O}(N_{\text{or}}^{-1})$ error scaling. The emulator is not a claim of wall-clock speedup on hardware, and we do not present it as one. The genuine, hardware-independent content is the query-complexity result of Lemma 2 and its incorporation into the end-to-end bound of

Theorem 1. These are mathematical facts about the algorithm; the emulator merely visualises them and lets us study the constants.

“Does distribution loading destroy the advantage?” Possibly, and we say so. The remark following Corollary 1 and the resource accounting in Table 6 are deliberately foregrounded. The advantage is cleanest for the repeated, online evaluation of many functionals against a structured distribution loaded once; it is weakest when an unstructured distribution must be reloaded for every query, in which case loading can be $\mathcal{O}(1/\varepsilon^2)$ -costly and dominate. This is the central caveat of the entire quantum-finance programme (Herman et al., 2023; Chakrabarti et al., 2021), and our modular bound makes it auditable rather than hiding it.

“Why a toy income-fluctuation model rather than a full HANK model?” Because the contribution is methodological and the toy model admits a trusted ground truth against which every error term can be measured. The framework is agnostic to the model: stage 1 can be any solver, and stage 2 any surrogate. Scaling the identical pipeline to a HANK model (Kaplan et al., 2018) or a continuous-time mean-field model (Achdou et al., 2022) changes the dimension d and the cost of stages 1–2 but not the structure of the bound or the $\mathcal{O}(N_{\text{or}}^{-1})$ estimation rate.

“Is the novelty sufficient?” We do not claim a new model or primitive. The novelty is the explicit, modular, end-to-end error accounting that unifies the solution and estimation stages and delineates the regime of advantage—an accounting that, to our reading, is implicit but not assembled in this form in the closest prior work (Skavysh et al., 2023). We position the paper as a refinement and consolidation, which is the honest description.

13.2 Limitations

Three limitations bound the contribution. First, the hardware caveat: canonical amplitude estimation needs deep circuits, and even iterative variants need qubit counts and gate fidelities beyond robust current capability for high-precision macro-financial functionals (Preskill, 2018; Bank for International Settlements, 2024). Second, the loading bottleneck discussed above. Third, our numerical evidence is an emulation; a hardware demonstration on a small instance, in the spirit of Stamatopoulos et al. (2020), is the natural next step but is outside the scope of a methodological paper. None of these undercuts the analytical results; they bound their practical reach today.

13.3 Policy relevance

For central banks and supervisors, the relevance is prospective. Stress testing is exactly the “repeated evaluation of many tail functionals against a structured scenario distribution” workload for which the advantage of Corollary 1 is cleanest, which is why policy institutions are tracking the technology (Gorbanyov et al., 2021; Bank for International Settlements, 2024; Hull et al., 2020). The contribution of this paper to that conversation is a clear, modular framework for reasoning about *when* the advantage is real and *how large* it is, together with a reproducible benchmark for measuring it.

13.4 From NISQ to fault tolerance

The resource accounting of Table 6 is a NISQ-era statement: it counts logical operations as if they were physical. On error-corrected hardware the picture changes in two offsetting ways. Fault tolerance removes the depth ceiling that today forces the iterative variant, so canonical amplitude estimation’s single deep circuit becomes admissible; but each logical Grover application then costs many physical operations through the error-correcting code, inflating the constant. The threshold analyses of [Chakrabarti et al. \(2021\)](#) for derivative pricing suggest that a practically useful advantage for financial functionals is a fault-tolerant proposition, not a NISQ one. Our framework is agnostic to this transition: the query-complexity results and the error decomposition hold in both regimes; only the mapping from oracle calls to wall-clock time changes.

13.5 Cryptographic considerations

A complete assessment of quantum computing for the financial system must weigh the opportunity studied here against the cryptographic risk that the same technology poses to the cryptographic primitives underpinning financial infrastructure, a theme emphasised by [Gorbanyov et al. \(2021\)](#) and [Bank for International Settlements \(2024\)](#). These are distinct channels—one computational, one security—and our paper speaks only to the first. We note the second for completeness and to avoid the impression that the prospective benefits can be assessed in isolation from the systemic risks.

13.6 A research agenda

Three lines of work would sharpen the contribution. First, structured loading: identifying the families of scenario distributions (log-concave, low-rank, copula-structured) for which the loading circuit is efficient, so that the advantage of Corollary 1 is preserved end to end. Second, a small-instance hardware demonstration of the baseline stress functional, in the spirit of [Stamatopoulos et al. \(2020\)](#), to move from emulated query complexity to measured behaviour on a device. Third, deployment of the identical pipeline to a solved HANK-plus-banking model (Section 12), where the solution stage is harder but the estimation stage and its bound are unchanged. Each is a concrete, falsifiable step rather than a promissory note.

14 Comparison with Classical Alternatives

Before concluding, we situate the framework against the classical methods a practitioner would otherwise reach for, to clarify what is and is not gained.

14.1 Quasi-Monte Carlo

Quasi-Monte Carlo (QMC) replaces pseudo-random points by low-discrepancy sequences and, for integrands of bounded variation in the Hardy–Krause sense, attains an error of $\mathcal{O}((\log N)^d/N)$, nearly $\mathcal{O}(N^{-1})$ in favourable cases ([Caffisch, 1998](#)). This is, asymptotically, of the same order as amplitude estimation’s $\mathcal{O}(N_{\text{or}}^{-1})$. Two caveats temper the comparison. First, the QMC rate

degrades with dimension through the $(\log N)^d$ factor and is fragile for the discontinuous indicator integrands of tail-probability estimation, where the Hardy–Krause variation is large or infinite. Amplitude estimation’s rate, by contrast, is $\mathcal{O}(N_{\text{or}}^{-1})$ regardless of the smoothness of f , because it estimates an amplitude rather than integrating a function. Second, the two are not mutually exclusive: QMC-style variance reduction can be composed with amplitude estimation, as Proposition 2 notes, improving the constant. We therefore view QMC as a classical benchmark that amplitude estimation matches in rate and can incorporate, rather than as a competitor that dominates it.

14.2 Importance sampling for rare events

For deep-tail functionals, importance sampling (IS) is the classical method of choice: it tilts the sampling measure toward the rare event, reducing variance by orders of magnitude. IS attacks the *constant* in the $\mathcal{O}(N^{-1/2})$ rate, not the rate itself. Amplitude estimation attacks the *rate*. The two are again complements: an importance-tilted distribution can be loaded as the state-preparation measure, combining the variance reduction of IS with the quadratic rate improvement of amplitude estimation. The fixed-relative-accuracy analysis of Corollary 1 is the natural language for this combination, since IS is precisely a device for controlling relative error in the tail.

14.3 The quantum-annealing route

A distinct quantum approach solves the dynamic program directly by encoding the Bellman optimisation as an energy-minimisation problem on a quantum annealer (Fernández-Villaverde and Hull, 2023). That route targets the *solution* stage, whereas QMC-Macro keeps the solution classical (or deep-learning-based) and targets the *estimation* stage. The two are addressed at different points of the pipeline and could in principle be combined: an annealer-produced solution feeding a gate-model amplitude-estimation read-out. We do not pursue the combination here, but note that the modular architecture of Section 6 would accommodate it by swapping stage 1.

14.4 Summary of the comparison

The recurring theme is composition rather than competition. Deep learning breaks the curse at the solution stage; classical variance reduction lowers the estimation constant; amplitude estimation improves the estimation rate; and annealing offers an alternative solution route. QMC-Macro’s contribution is to assemble the deep-learning solution and the amplitude-estimation read-out into one auditable pipeline and to characterise the regime in which the composite advantage is real. It does not claim to dominate any single classical method on that method’s own axis.

15 Conclusion

We have developed QMC-Macro, a unified framework that couples a deep-learning surrogate for high-dimensional macroeconomic policy functions with quantum amplitude estimation for the expectation functionals that drive policy evaluation and stress testing. The framework is offered not as a new model or a new algorithm but as a careful consolidation and extension of an existing literature, with an explicit end-to-end error decomposition (Theorem 1) that separates the surrogate, loading, and estimation contributions and makes the regime of quantum advantage auditable (Corollary 1). Our reproducible numerical study shows that the surrogate reproduces the optimal policy to within half a percent and that a faithful emulator of iterative amplitude estimation attains the theoretical $\mathcal{O}(N_{\text{or}}^{-1})$ rate, implying sample-complexity speedups that grow with the required precision. We have been deliberate about the caveats: the loading bottleneck, the NISQ resource costs, and the emulated nature of the hardware evidence. The natural extensions are a hardware demonstration on a small instance, the integration of structured loading circuits that preserve the advantage, and the deployment of the identical pipeline to a full HANK or financial-sector model, where the solution stage is harder but the estimation stage is unchanged.

A Proofs and Additional Derivations

A.1 Derivation of the Euler equation

We derive (5) from the Bellman equation (1). Substituting the budget constraint (2) to eliminate $c = (1+r)a + wz - a'$, the household chooses a' to maximise $u((1+r)a + wz - a') + \beta \mathbb{E}[V(a', z') \mid z]$. The first-order condition with respect to a' , accounting for the constraint $a' \geq \underline{a}$ with multiplier $\lambda \geq 0$, is

$$-u'(c) + \beta \mathbb{E}[V_a(a', z') \mid z] + \lambda = 0, \quad \lambda(a' - \underline{a}) = 0.$$

The envelope theorem applied to (1) gives $V_a(a, z) = (1+r)u'(c(a, z))$. Advancing one period, $V_a(a', z') = (1+r)u'(c')$, and substituting yields $u'(c) = \beta(1+r) \mathbb{E}[u'(c') \mid z] + \lambda$, which is (5). When the constraint does not bind $\lambda = 0$ and the equation holds with equality; when it binds $a' = \underline{a}$ and $\lambda > 0$ absorbs the slack. ■

A.2 Contraction property of policy iteration

Lemma 5 (Geometric convergence of EGM). *Under $\beta(1+r) < 1$ and CRRA utility, the policy-iteration operator implied by Algorithm 1 is a contraction in the sup-norm on consumption policies, with modulus bounded by a constant $\beta(1+r) < 1$, so the iterates converge geometrically to the unique fixed point.*

Proof sketch. The Coleman–Reffett policy operator T defined by the Euler equation is monotone and discounts: for two candidate consumption functions $c_1 \leq c_2$, monotonicity of u' and the conditional expectation give $Tc_1 \leq Tc_2$, and scaling consumption by $1 + \delta$ scales the Euler right-hand side sub-proportionally by the factor $\beta(1+r)$ under CRRA, establishing Blackwell's

discounting condition with modulus $\beta(1+r)$. By Blackwell's sufficient conditions T is a contraction of modulus $\beta(1+r) < 1$, and the contraction mapping theorem delivers a unique fixed point with geometric convergence. The endogenous-grid implementation computes T exactly on the grid up to interpolation error, inheriting the rate; the observed 333 iterations to tolerance 10^{-9} are consistent with modulus $\beta(1+r) = 0.96 \times 1.03 \approx 0.989$, since $0.989^{333} \approx 0.025 \cdot 10^{-3}$. $\square$

A.3 Full proof of the amplitude-estimation geometry

We expand the proof sketch of Lemma 2. Write $|\Psi\rangle = \mathcal{A}|0\rangle = \cos\theta_a|\Psi_0\rangle + \sin\theta_a|\Psi_1\rangle$, where $|\Psi_1\rangle = |\psi_1\rangle|1\rangle$ is the (normalised) marked component and $|\Psi_0\rangle = |\psi_0\rangle|0\rangle$ the unmarked one, with $a = \sin^2\theta_a$. The reflection $S_\chi = I - 2|\Psi_1\rangle\langle\Psi_1|$ negates the marked component; $S_0 = I - 2|0\rangle\langle 0|$ reflects about the initial state, so $\mathcal{A}S_0\mathcal{A}^{-1} = I - 2|\Psi\rangle\langle\Psi|$ reflects about $|\Psi\rangle$. The product $\mathcal{Q} = (I - 2|\Psi\rangle\langle\Psi|)(I - 2|\Psi_1\rangle\langle\Psi_1|)$ is, on the real two-dimensional span of $\{|\Psi_0\rangle, |\Psi_1\rangle\}$, the composition of two reflections, hence a rotation; the angle between the two reflection axes is θ_a , so the rotation angle is $2\theta_a$. Thus the eigenvalues of $\mathcal{Q}$ restricted to this plane are $e^{\pm 2i\theta_a}$. Quantum phase estimation with a control register of t qubits, i.e. $M = 2^t - 1$ applications of $\mathcal{Q}$, returns an estimate $\hat{\theta}_a$ with $|\hat{\theta}_a - \theta_a| \leq \pi/M$ with probability at least $8/\pi^2$ (Brassard et al., 2002). Finally, $|\hat{a} - a| = |\sin^2\hat{\theta}_a - \sin^2\theta_a| = |\sin(\hat{\theta}_a + \theta_a)\sin(\hat{\theta}_a - \theta_a)| \leq |\sin(\hat{\theta}_a - \theta_a)| \leq 2\sqrt{a(1-a)}(\pi/M) + (\pi/M)^2$ after bounding $|\sin(\hat{\theta}_a + \theta_a)| \leq 1$ in the second-order term and using $|\sin 2\theta_a| = 2\sqrt{a(1-a)}$ in the first, giving (15). $\blacksquare$

A.4 Loading discretisation error versus qubit count

Lemma 6 (Discretisation error of loading). *Suppose X is supported on a bounded interval discretised into 2^n equal cells, and the loaded distribution assigns each cell its true probability mass. If the loss L is Lipschitz with constant κ_L and the density of X is bounded, the total-variation loading error from discretisation satisfies $\varepsilon_{\text{load}} = \mathcal{O}(2^{-n})$, so achieving $\varepsilon_{\text{load}} \leq \varepsilon$ requires $n = \mathcal{O}(\log(1/\varepsilon))$ value qubits.*

Proof. Equal-cell discretisation of a bounded-density distribution incurs total-variation error proportional to the cell width $\mathcal{O}(2^{-n})$ by a standard quadrature argument; the Lipschitz loss controls how this propagates to the comparator outcome. Inverting $2^{-n} \leq \varepsilon$ gives the qubit count. $\square$

This is the benign case. The difficulty flagged after Corollary 1 is that *preparing* the amplitudes $\sqrt{p_x}$ by a quantum circuit, as opposed to merely choosing the cells, can require circuit depth growing with 2^n for unstructured p ; structured families (e.g. log-concave or low-rank) admit efficient loading, which is the regime in which the overall advantage is preserved.

A.5 Discontinuous integrands

The surrogate-error bound in Theorem 1 uses $\sup_x |f_\theta(x) - f(x)|$, which is loose for the indicator functional $f = \mathbf{1}\{L > \ell\}$ because a small perturbation of the policy can flip f on a set of states near the threshold. A tighter control uses the $L^1(\mathbb{P}_X)$ norm. If $L(X)$ has a density bounded

by $\bar{p}$ in a neighbourhood of ℓ and the surrogate perturbs the loss by at most $\kappa_L \varepsilon_{\text{NN}}$ pointwise, then

$$\mathbb{E}_{\mathbb{P}_X} |f_\theta - f| = \mathbb{P}(|L(X) - \ell| \leq \kappa_L \varepsilon_{\text{NN}}) \leq 2\bar{p} \kappa_L \varepsilon_{\text{NN}},$$

so the surrogate term in (16) can be replaced by $2\bar{p} \kappa_L \varepsilon_{\text{NN}}$, which is linear in the surrogate error with a constant governed by the density of the loss at the threshold. This refinement matters precisely in the deep tail, where $\bar{p}$ is small and the surrogate term is correspondingly forgiving.

A.6 Optimal split of the error budget

Given a fixed total budget for the three stages, the bound (16) is minimised by equalising marginal costs across stages. Since the estimation term falls as $\mathcal{O}(N_{\text{or}}^{-1})$ in oracle calls while the surrogate and loading terms fall at stage-specific (one-off) rates, the optimal allocation drives the one-off terms below the target ε and then spends the online budget on N_{or} . This is the formal content of the “one-off pre-processing” hypothesis of Corollary 1.

A.7 The Rouwenhorst construction

For completeness we record the Rouwenhorst (1995) discretisation of (4). Set $p = q = (1 + \rho_z)/2$. Define $\Theta_2 = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$ and build Θ_n from Θ_{n-1} by the recursion

$$\Theta_n = p \begin{bmatrix} \Theta_{n-1} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} + (1-p) \begin{bmatrix} \mathbf{0} & \Theta_{n-1} \\ 0 & \mathbf{0}^\top \end{bmatrix} + (1-q) \begin{bmatrix} \mathbf{0}^\top & 0 \\ \Theta_{n-1} & \mathbf{0} \end{bmatrix} + q \begin{bmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \Theta_{n-1} \end{bmatrix},$$

normalising all but the first and last rows to sum to one. The state grid is equally spaced on $[-\psi, \psi]$ with $\psi = \sigma_z \sqrt{(n_z - 1)/(1 - \rho_z^2)}$. The chain matches the persistence ρ_z and the unconditional variance $\sigma_z^2/(1 - \rho_z^2)$ of the continuous process exactly, which is the property that makes it accurate for the highly persistent income processes used in the wealth-distribution literature.

A.8 From tail probabilities to expected shortfall

The read-out stage converts the estimated exceedance probabilities into the risk measures supervisors report. Value-at-risk at level α is the threshold ℓ_α solving $\mathbb{P}(L > \ell_\alpha) = 1 - \alpha$, obtained by a monotone sweep of the comparator threshold and inversion of the estimated curve $\ell \mapsto \hat{a}(\ell)$. Expected shortfall is the conditional tail mean,

$$\text{ES}_\alpha = \mathbb{E}[L \mid L > \ell_\alpha] = \ell_\alpha + \frac{1}{1 - \alpha} \int_{\ell_\alpha}^{\infty} \mathbb{P}(L > \ell) d\ell,$$

so a discretised sweep of exceedance probabilities $\{\hat{a}(\ell_k)\}$ yields ES_α by numerical integration. Each $\hat{a}(\ell_k)$ is itself an amplitude estimate, so the entire risk-measure pipeline inherits the $\mathcal{O}(N_{\text{or}}^{-1})$ accuracy of the estimation stage, amortising the one-off loading cost across the whole sweep—precisely the repeated-functional workload for which the advantage is cleanest.

A.9 IQAE confidence intervals

The certified interval reported in the read-out is constructed as follows (Grinko et al., 2021). In each round k , the estimator measures the objective qubit after m_k Grover iterations a fixed number of times, forming a frequency estimate of $\sin^2((2m_k + 1)\theta_a)$. A Chernoff–Hoeffding bound converts this frequency into a confidence interval for $(2m_k + 1)\theta_a \bmod \pi$; the schedule chooses m_k so that this interval, mapped back to θ_a , is unambiguous (does not straddle a half-period boundary). Intersecting the per-round intervals contracts the estimate geometrically, and a union bound over the $\mathcal{O}(\log(1/\varepsilon))$ rounds controls the overall coverage at the nominal level. The output is a two-sided interval for $a = \sin^2 \theta_a$ with guaranteed coverage, which is exactly the object a supervisor needs to report a stress-test exceedance probability with a stated confidence level.

B Algorithms

Algorithm 2. QMC-Macro stress-test estimation (high level)

1. **Solve.** Solve the macroeconomic model (1)–(5) by the endogenous grid method to obtain ground-truth policy $c(\cdot)$ and value $V(\cdot)$; verify via Euler residuals.
 2. **Train.** Fit the surrogate $\hat{c}_\theta$ in (10) by minimising the loss (11) with Adam; validate against the ground truth (η_{NN}).
 3. **Load.** Construct $\mathcal{A}$: load the discretised scenario distribution $\mathbb{P}_X^{(n)}(\varepsilon_{\text{load}})$ and rotate the objective qubit by the comparator encoding $f(x) = \mathbf{1}\{L(x) > \ell\}$, where L is built from $\hat{c}_\theta$.
 4. **Estimate.** Run iterative amplitude estimation on $\mathcal{Q} = \mathcal{A}S_0\mathcal{A}^{-1}S_X$ with N_{or} oracle calls to obtain $\hat{a}$ and a certified confidence interval.
 5. **Read out.** Report $\hat{a} = \mathbb{P}(L(X) > \ell)$; sweep ℓ for value-at-risk and integrate for expected shortfall.
-

C Data Appendix

The macro-financial panel (2000–2026) underlying the scenario calibration is provided in the companion workbook with one sheet of documented sources (variable, primary source, series identifier, database) and one sheet of regime-conditional shock moments. Table 8 reports the shock moments used to define the five scenarios of Table 5; the means and standard deviations are expressed in the standardised loss units used by the risk functional.

Table 8: Shock moments by regime (standardised loss units).

Regime	Mean loss	Std. dev.	Threshold
Baseline	-0.50	1.00	1.000
Adverse	-0.80	1.30	1.000
Severely adverse	-1.20	1.60	1.000
Tail (1%)	-0.50	1.00	2.326
Tail (0.1%)	-0.50	1.00	3.090

Table 9: Macro-financial panel, 2000–2026 (regime-calibrated representative series). Columns: GDP growth, CPI inflation, policy rate, credit-to-GDP, equity return, equity volatility, sovereign spread (bps), bank capital ratio, and non-performing-loan ratio. Crisis episodes are imposed for 2008–2009 and 2020. See the companion workbook’s **DataSources** sheet for live-vintage series identifiers.

Year	GDPg	CPI	Policy	Cr/GDP	EqRet	EqVol	Spread	BankCap	NPL
2000	2.21	3.82	3.69	90.10	-0.12	17.46	204.30	11.34	3.33
2001	3.24	1.44	2.53	91.10	9.18	19.44	153.90	11.30	3.07
2002	1.10	3.21	3.66	91.50	6.02	18.41	191.60	11.98	2.78
2003	4.40	1.14	2.18	93.10	15.19	22.86	80.30	11.95	3.47
2004	3.64	3.39	3.23	96.60	0.05	16.34	21.60	12.26	3.16
2005	2.71	3.25	3.36	96.50	1.82	17.84	180.70	12.55	3.12
2006	2.69	3.50	3.46	95.60	5.50	16.08	63.30	12.53	2.11
2007	3.30	1.70	2.62	94.50	2.07	20.99	123.60	12.96	2.12
2008	-3.27	4.93	5.23	93.60	-40.39	47.31	478.70	13.18	3.22
2009	-3.23	3.95	4.83	91.70	-5.99	46.29	432.00	13.44	7.05
2010	5.72	1.32	1.98	90.90	0.50	18.84	182.10	13.46	7.08
2011	3.51	2.75	3.00	94.50	-4.91	17.38	532.00	13.87	3.10
2012	2.94	3.22	3.30	94.30	-2.43	19.36	492.10	13.85	3.25
2013	2.91	2.69	3.09	95.80	8.67	17.38	118.80	13.64	3.02
2014	3.16	1.95	2.75	94.40	24.75	18.04	127.70	14.08	3.09
2015	2.39	2.02	2.93	95.70	3.11	19.12	76.20	14.14	3.15
2016	2.60	3.78	3.59	93.30	8.81	18.16	110.40	14.30	3.56
2017	3.55	2.81	3.01	92.90	1.15	20.17	115.80	14.83	2.69
2018	2.87	3.11	3.27	98.40	-0.76	16.64	85.90	15.21	3.47
2019	1.63	4.66	4.14	98.20	-11.77	20.87	154.00	15.49	3.38
2020	-3.48	3.79	4.81	101.40	-23.67	44.85	412.90	15.85	7.06
2021	5.66	3.75	2.97	103.40	23.56	19.18	101.90	15.86	6.71
2022	2.77	9.09	5.68	104.00	4.46	22.11	29.90	16.16	3.22
2023	2.85	3.84	3.57	103.70	18.72	19.30	139.90	16.53	3.52
2024	3.64	2.97	3.06	107.30	11.50	18.46	149.70	16.83	2.36
2025	4.82	1.42	2.20	108.00	13.17	20.39	144.90	16.94	3.04
2026	2.29	3.51	3.55	112.00	10.01	17.36	236.70	16.34	3.04

D Reproducibility

All numerical results are generated by the accompanying Python code under random seed 20260604. The pipeline comprises: `engine.py` (model, classical Monte Carlo, amplitude-estimation emulator, surrogate); `run_experiments.py` (experiment suite, writing `results.json`); `make_figures.py` (all figures at 300 dpi); and the workbook generators for the data, estimation, and annex spreadsheets. The bibliography is provided in both BibTeX and RIS formats for direct import into reference managers such as Zotero.

E Glossary of Quantum Terms

For the convenience of the economist reader, we collect the quantum-computational terms used in the paper.

Qubit. A two-level quantum system; a unit vector $\alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2$. Measurement returns 0 or 1 with probabilities $|\alpha|^2, |\beta|^2$.

Superposition. A linear combination of basis states; an n -qubit register holds a superposition over 2^n basis states simultaneously.

Amplitude. A complex coefficient α_x in a superposition; the probability of outcome x is $|\alpha_x|^2$. The target $a = \mathbb{E}[f(X)]$ is encoded as the squared amplitude of a marked state.

Gate. A unitary operation on qubits (e.g. Hadamard, controlled-NOT). Circuits compose gates.

Circuit depth. The number of sequential gate layers; the binding resource on NISQ hardware because of decoherence.

Grover operator $\mathcal{Q}$. $\mathcal{A}S_0\mathcal{A}^{-1}S_X$; rotates the state by $2\theta_a$ in the good/bad plane, with $a = \sin^2 \theta_a$.

Amplitude estimation (AE). A family of algorithms estimating a with error $\mathcal{O}(N_{\text{or}}^{-1})$ in the number N_{or} of Grover applications; a quadratic improvement over classical sampling.

Quantum phase estimation (QPE). The subroutine used by canonical AE to read off the rotation angle; requires deep circuits.

Iterative AE (IQAE). A phase-estimation-free variant with bounded per-round depth and certified confidence intervals (Grinko et al., 2021); our reference estimator.

State preparation / distribution loading. The unitary that loads a discretised probability distribution onto qubits; the practical bottleneck for unstructured distributions.

NISQ. Noisy intermediate-scale quantum hardware (Preskill, 2018): limited qubits and depth.

Oracle / Grover call. One application of $\mathcal{Q}$; the unit in which query complexity is measured.

F Notation

Table 10: Principal notation.

Symbol	Meaning
a, z	Individual assets and (idiosyncratic) productivity
V, c, g	Value function, consumption policy, savings policy
β, γ, r, w	Discount factor, risk aversion, interest rate, wage
Π	Productivity transition matrix (Rouwenhorst)
μ	Cross-sectional distribution of (a, z)
X	Aggregate risk-factor vector, $X \in \mathbb{R}^d$
f, L, ℓ	Integrand, loss function, threshold
$a = \mathbb{E}[f(X)]$	Target amplitude / tail probability
$\hat{c}_\theta$	Neural surrogate with parameters θ
$\mathcal{A}, \mathcal{Q}$	State-preparation and Grover operators
N_{or}	Number of Grover (oracle) applications
ε	Target accuracy
$\eta_{\text{NN}}, \varepsilon_{\text{load}}$	Surrogate and loading error terms
θ_a	Amplitude angle, $a = \sin^2 \theta_a$
d	Dimension of the aggregate state

References

- Aboussalah, A. M., Chi, C., and Lee, C.-G. (2023). Quantum computing reduces systemic risk in financial networks. *Scientific Reports*, 13(1):3839.
- Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2022). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies*, 89(1):45–86.
- Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684.
- An, D., Linden, N., Liu, J.-P., Montanaro, A., Shao, C., and Wang, J. (2021). Quantum-accelerated multilevel monte carlo methods for stochastic differential equations in mathematical finance. *Quantum*, 5:481.
- Azinovic, M., Gaegauf, L., and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, 63(4):1471–1525.
- Bank for International Settlements (2024). Quantum computing and the financial system: opportunities and risks. BIS Papers 149, Bank for International Settlements.
- Bouland, A., van Dam, W., Joorati, H., Kerenidis, I., and Prakash, A. (2020). Prospects and challenges of quantum finance. Technical Report 2011.06492, arXiv.
- Brassard, G., Høyer, P., Mosca, M., and Tapp, A. (2002). Quantum amplitude amplification and estimation. In *Quantum Computation and Information*, volume 305 of *Contemporary Mathematics*, pages 53–74. American Mathematical Society.

- Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, 104(2):379–421.
- Cafisch, R. E. (1998). Monte carlo and quasi-monte carlo methods. *Acta Numerica*, 7:1–49.
- Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320.
- Chakrabarti, S., Krishnakumar, R., Mazzola, G., Stamatopoulos, N., Woerner, S., and Zeng, W. J. (2021). A threshold for quantum advantage in derivative pricing. *Quantum*, 5:463.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314.
- Egger, D. J., Gutiérrez, R. G., Mestre, J. C., and Woerner, S. (2020). Credit risk analysis using quantum computers. *IEEE Transactions on Computers*, 70(12):2136–2145.
- Fernández-Villaverde, J. and Hull, I. J. (2023). Dynamic programming in economics on a quantum annealer. NBER Working Paper 31326, National Bureau of Economic Research.
- Fernández-Villaverde, J., Nuño, G., and Perla, J. (2024). Taming the curse of dimensionality: Quantitative economics with deep learning. NBER Working Paper 33117, National Bureau of Economic Research.
- Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer.
- Gorbanyov, M., Malaika, M., and Saadi Sedik, T. (2021). Quantum computing and the financial system: Spooky action at a distance? IMF Working Paper 2021/071, International Monetary Fund.
- Grinko, D., Gacon, J., Zoufal, C., and Woerner, S. (2021). Iterative quantum amplitude estimation. *npj Quantum Information*, 7(1):52.
- Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)*, pages 212–219.
- Gu, Z., Laurière, M., Merkel, S., and Payne, J. (2024). Global solutions to master equations for continuous time heterogeneous agent macroeconomic models. Technical Report 2406.13726, arXiv.
- Han, J., Jentzen, A., and E, W. (2018). Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510.
- Han, J., Yang, Y., and E, W. (2021). Deepham: A global solution method for heterogeneous agent models with aggregate shocks. Technical Report 2112.14377, arXiv.
- Heinrich, S. (2002). Quantum summation with an application to integration. *Journal of Complexity*, 18(1):1–50.

- Herman, D., Googin, C., Liu, X., Galda, A., Safro, I., Sun, Y., Pistoia, M., and Alexeev, Y. (2023). Quantum computing for finance. *Nature Reviews Physics*, 5:450–465.
- Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.
- Hull, I., Sattath, O., Diamanti, E., and Wendin, G. (2020). Quantum technology for economists. Working Paper 398, Sveriges Riksbank.
- Judd, K. L. (1998). *Numerical Methods in Economics*. MIT Press.
- Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, 108(3):697–743.
- Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.
- Krusell, P. and Smith, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896.
- Maliar, L., Maliar, S., and Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics*, 122:76–101.
- Montanaro, A. (2015). Quantum speedup of monte carlo methods. *Proceedings of the Royal Society A*, 471(2181):20150301.
- Nielsen, M. A. and Chuang, I. L. (2010). *Quantum Computation and Quantum Information*. Cambridge University Press, 10th anniversary edition.
- Orús, R., Mugel, S., and Lizaso, E. (2019). Forecasting financial crashes with quantum computing. *Physical Review A*, 99(6):060301.
- Preskill, J. (2018). Quantum computing in the nisq era and beyond. *Quantum*, 2:79.
- Rebentrost, P., Gupt, B., and Bromley, T. R. (2018). Quantum computational finance: Monte carlo pricing of financial derivatives. *Physical Review A*, 98(2):022321.
- Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium business cycle models. In Cooley, T. F., editor, *Frontiers of Business Cycle Research*, pages 294–330. Princeton University Press.
- Sirignano, J. and Spiliopoulos, K. (2018). Dgm: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375:1339–1364.
- Skavysh, V., Priazhkina, S., Guala, D., and Bromley, T. R. (2023). Quantum monte carlo for economics: Stress testing and macroeconomic deep learning. *Journal of Economic Dynamics and Control*, 153:104680.
- Stamatopoulos, N., Egger, D. J., Sun, Y., Zoufal, C., Iten, R., Shen, N., and Woerner, S. (2020). Option pricing using quantum computers. *Quantum*, 4:291.

- Suzuki, Y., Uno, S., Raymond, R., Tanaka, T., Onodera, T., and Yamamoto, N. (2020). Amplitude estimation without phase estimation. *Quantum Information Processing*, 19(2):75.
- Woerner, S. and Egger, D. J. (2019). Quantum risk analysis. *npj Quantum Information*, 5(1):15.