Quantum Econometrics: How to Explain Its Quantitative Successes and How the Resulting Formulas Are Related to Scale Invariance, Entropy, and Fuzziness

Kittawit Autchariyapanitkul[1], Olga Kosheleva[2]
Vladik Kreinovich[2], Songsak Sriboonchitta[3]

[1]Faculty of Economics, Maijo University
Chiang Mai, Thailand, kittawit_a@mju.ac.th
[2]University of Texas at El Paso, El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu
[3]Faculty of Economics, Chiang Mai University
Chiang Mai 50200 Thailand, songsajecon@gmail.com

1. Formulation of the Problem

- Many aspects of human behavior seem to be well-described by formulas from quantum physics.
- This is understandable on the qualitative level: similar to quantum physics,
  - every time we gain new knowledge,
  - we inevitably change the system.
- For example:
  - once we learn a new dependence between the economic variables,
  - we can make better predictions of economic phenomena and thus,
  - change the behavior of decision makers.
- But how can we explain this success on the quantitative level?

2. Non-quantum vs. Quantum Probabilities: 2-Slot Experiment

- We have a particle generator – e.g., a light source or a radio source that generates photons.
- There is an array of sensors at different distances from this generator.
- By detecting the particles, we estimate the probability density $\rho(x)$ corresponding to the sensor's location x.
- There is a barrier with 2 slots between the generator and the sensors.
- If we open only the 1st slot, we get $\rho_1(x)$.
- If we open only the 2nd slot, we get $\rho_2(x)$.
- If we open both slots, hat will then be the resulting probability density $\rho(x)$?

3. Non-quantum vs. Quantum (cont-d)

- A particle reaches the sensor if it either went through the 1st slot or through the 2nd slot.
- THus, in non-quantum physics, $\rho(x) = \rho_1(x) + \rho_2(x)$.
- However, this is *not* what we observe.
- In quantum physics, we need a complex-valued function $\psi(x)$ – called *wave function* – for which
$$\rho(x) = |\psi(x)|^2.$$
- In 2-slot case, $\psi(x) = \psi_1(x) + \psi_2(x)$.
- If both $\rho_i(x)$ are positive real numbers, we have
$$\rho(x) = \left(\sqrt{\rho_1(x)} + \sqrt{\rho_2(x)}\right)^2 \neq \rho_1(x) + \rho_2(x).$$
- In general,
$$\left(\sqrt{\rho_1(x)} - \sqrt{\rho_2(x)}\right)^2 \le \rho(x) \le \left(\sqrt{\rho_1(x)} + \sqrt{\rho_2(x)}\right)^2.$$

4. How is This Related to Human Behavior

- In the early 1980s, researchers from the Republic of Georgia observed kids in a 2-door room.
- On the other side, boxes with treats were placed.
- In some cases, both doors were open, in other cases, only one door was open.
- researchers measured how frequently kids pick up treats at different location x.
- For the older kids, $f(x) \approx f_1(x) + f_2(x)$.
- However, for younger kids (3-4 years old), they got:

$$f(x) \approx \left(\sqrt{f_1(x)} + \sqrt{f_2(x)}\right)^2.$$

- Some aspects of adult behavior are also well described by quantum formulas.
- This is the main idea behind quantum econometrics.

5. Analysis of the Problem

- Based on $\rho_1(x)$ and $\rho_2(x)$, we want to estimate $\rho(x)$:
$$\rho(x) \approx f(\rho_1(x), \rho_2(x)).$$
- What are the natural properties of the corresponding algorithm $f(a, b)$?
- It does not matter which door is called st, so we should have $f(a, b) = f(b, a)$ – *commutativity*.
- Small changes in $\rho_i(x)$ should lead to small changes in $\rho(x)$, so $f(a, b)$ should be *continuous*.
- If $\rho_i(x)$ increases, $\rho(x)$ should increase, so $f(a, b)$ should be *monotonic*.
- For 3 slots, the estimate should be the same whether we first combine 1-2 or 2-3:
$$f(f(a, b), c) = f(a, f(b, c)) \text{ – } \textit{associativity}.$$

6. Final Natural Requirement: Scale-Invariance

- By definition, the probability density is probability divided by the length or area (or volume).
- We can use different units for measuring length, and thus, different units for measuring area or volume:
  - if we replace the original measuring unit with the one which is λ times smaller,
  - the numerical values of the probability density gets multiplied by λ.
- It is reasonable to require that the estimating function $f(a,b)$ should not change after this re-scaling:
  - if $\rho(x)=f(\rho_1(x),\rho_2(x))$
  - then $\lambda\cdot\rho(x)=f(\lambda\cdot\rho_1(x),\lambda\cdot\rho_2(x))$.
- Thus, we require that $f(\lambda\cdot a,\lambda\cdot b)=\lambda\cdot f(a,b)$ for all possible values a, b, and λ.

7. Main Result

Definition. *We say that* $f : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$ *is a* scale-invariant estimation function *if it is:*

- *commutative, associative, continuous,*
- *(non-strictly) increasing,*
- *and* $f(\lambda \cdot a, \lambda \cdot b) = \lambda \cdot f(a, b)$ *for all* a*,* b*, and* λ*.*

Proposition. *The only scale-invariant estimation functions are:*

$$f(a, b) = 0, \quad f(a, b) = \min(a, b), \quad f(a, b) = \max(a, b),$$

$$\text{and } f(a, b) = (a^\alpha + b^\alpha)^{1/\alpha} \text{ for some } \alpha.$$

8. Discussion

- *Reminder:* $f(a,b) = (a^\alpha + b^\alpha)^{1/\alpha}$.
- For $\alpha = 1$, we get the usual formula for the probability for the event $A \vee B$ when A and B are disjoint.
- For $\alpha = 0.5$, we get the quantum formula.
- Thus, we get the desired justified general formula for which special cases are:
  - the traditional probabilistic formula $f(a,b) = a + b$ for which
  $$\rho(x) = \rho_1(x) + \rho_2(x) \text{ and}$$
  - the quantum formula $f(a,b) = (a^{0.5} + b^{0.5})^2$, for which
  $$\rho(x) = \left(\sqrt{\rho_1(x)} + \sqrt{\rho_2(x)}\right)^2.$$

9. Relation to Entropy

- *Reminder:* $f(a,b)=(a^\alpha+b^\alpha)^{1/\alpha}$.
- Probabilistic case is $\alpha=1$, quantum is $\alpha=0.5$.
- Most aspects of human behavior can be described by the usual probabilistic formulas.
- This means that for human behavior, we have $\alpha\approx 1$.
- Quantum formulas capture some aspects of human behavior.
- Since quantum formulas correspond to $\alpha<1$, the actual α is $\alpha=1-\varepsilon$ for some small $\varepsilon>0$.
- In this case, $\rho(x)\approx\rho_1(x)+\rho_2(x)+\varepsilon\cdot\Delta\rho(x)$, where

$$\Delta\rho(x)\stackrel{\text{def}}{=}-\rho_1(x)\cdot\ln(\rho_1(x))-\rho_2(x)\cdot\ln(\rho_2(x))-$$
$$(-(\rho_1(x)+\rho_2(x))\cdot\ln(\rho_1(x)+\rho_2(x))).$$

- The sum of $-\rho(x)\cdot\ln(\rho(x))$ is Shannon's entropy!

10. Relation to Fuzziness

- For probabilities, $P(A \vee B) \leq P(A) + P(B)$, so
$$\rho(x) \leq \rho_1(x) + \rho_2(x).$$
- However, in the quantum case,
$$\rho(x) = \rho_1(x) + \rho_2(x) + 2\sqrt{\rho_1(x) \cdot \rho_2(x)} > \rho_1(x) + \rho_2(x).$$
- So, quantum formulas cannot be interpreted in terms of the probabilities; same for all $\alpha < 1$.
- The fact that some aspects of human behavior cannot be described in probabilistic terms is well known, e.g.,:
  - people estimate the probability that a person X is a professional *and* a feminist as higher than
  - the probability that this person is a feminist.
- So, $P(A)$, $P(B)$, $P(A \vee B)$ are not probabilities, but degrees of certainty – as in fuzzy logic.

11. Probabilistic Interpretation

- Another option: take into account that $\rho_1(x)$ is the probability that:
  - the particle passed through the 1st slot *and*
  - did not pass through the 2nd slot:

$$\rho_1(x) = P(A_1) - P(A \,\&\, A_2), \;\; \rho_2(x) = P(A_2) - P(A \,\&\, A_2).$$

- So, to describe the probability $C(u,v) \stackrel{\text{def}}{=} P(A_1 \,\&\, A_2)$ in terms of $u \stackrel{\text{def}}{=} P(A_1)$ and $v \stackrel{\text{def}}{=} P(A_2)$, we have:

$$(u + v - C(u,v))^\alpha = (u - C(u,v))^\alpha + (v - C(u,v))^\alpha.$$

- In the quantum case, we have an explicit expression

$$C(u,v) = \frac{u + v \pm \sqrt{(u+v)^2 - 12u \cdot v}}{3}.$$

- This operation $C(u,v)$ is *not* associative.

12. Probabilistic and Fuzzy Interpretations

- In the probabilistic interpretation, we assume that the particle cannot go through both slots.
- We ended up by allowing a non-zero probability that the particle goes through both slots.
- This is exactly what fuzzy does:
  - instead of assuming that a person is either young or not young,
  - it takes into account that a person can be to some extent young and to some extent not young.

13. Acknowledgments

- We acknowledge support of Center of Excellence in Econometrics, Chiang Mai University, Thailand.
- This work was also supported in part by the US National Science Foundation grant HRD-1242122.

14. Proof: Case When $f(1,1)=1$ and $f(0,1)=0$

- For $b>0$, scale-invariance implies $f(b\cdot 0, b\cdot 1)=b\cdot 0$, i.e., $f(0,b)=0$.
- By taking $b\to 0$ and using continuity, we also get $f(0,0)=0$; thus, $f(0,b)=0$ for all b.
- By commutativity, $f(a,0)=0$ for all a.
- Let us prove, by contradiction, that $f(1,a)\le 1$ for all a.
- Indeed, if for some a, $b\stackrel{\text{def}}{=}f(1,a)>1$, then, due to associativity and $f(1,1)=1$:

$$f(1,b)=f(1,f(1,a))=f(f(1,1),a)=f(1,a)=b.$$

- Due to scale-invariance with $\lambda=b$, the equality $f(1,b)=b$ implies that $f(b,b^2)=b^2$; thus:

$$f(1,b^2)=f(1,f(b,b^2))=f(f(1,b),b^2)=f(b,b^2)=b^2.$$

15. Case $f(1,1) = 1$ and $f(0,1) = 0$ (cont-d)

- Similarly, from $f(1, b^2) = b^2$, we conclude that:
  - for $b^4 = (b^2)^2$, we have $f(1, b^4) = b^4$, and,
  - in general, that $f(1, b^{2^n}) = b^{2^n}$ for every n.
- Scale invariance with $\lambda = b^{-2^n}$ implies $f(b^{-2^n}, 1) = 1$.
- In the limit $n \to \infty$, we get $f(0,1) = 1$, which contradicts to our assumption $f(0,1) = 0$.
- This contradiction shows that indeed, $f(1,a) \le 1$.
- For $a \ge 1$, monotonicity implies $1 = f(1,1) \le f(1,a)$.
- So, $f(1,a) \le 1$ implies that $f(1,a) = 1$.
- If $0 < a' \le b'$, then for $r \stackrel{\text{def}}{=} \dfrac{b'}{a'} \ge 1$, scale-invariance with $\lambda = a'$ implies $a' \cdot f(1,r) = f(a' \cdot 1, a' \cdot r) = f(a', b')$.
- So, $f(a,b) = \min(a,b)$ for all a and b.

16. Proof: Case When $f(1,1)=1$ and $f(0,1)>0$

- Let us show that in this case, $f(0,0)=0$.
- Indeed, scale-invariance with $\lambda=2$ implies that from $f(0,0)=a$, we can conclude that

$$f(2\cdot 0, 2\cdot 0)=f(0,0)=2\cdot a.$$

- Thus $a=2\cdot a$, hence $a=0$.
- Let us now prove that in this subcase, $f(0,1)=1$.
- Indeed, in this case, for $a\stackrel{\text{def}}{=}f(0,1)$, we have, due to $f(0,0)=0$ and associativity, that

$$f(0,a)=f(0,f(0,1))=f(f(0,0),1)=f(0,1)=a.$$

- Here, $a>0$, so by applying scale invariance with $\lambda=a^{-1}$, we conclude that $f(0,1)=1$.
- Let us prove that for every $a\le b$, we have $f(a,b)=b$.

17. Case $f(1,1) = 1$ and $f(0,1) > 0$ (cont-d)

- Let us prove that for every $a \le b$, we have $f(a,b) = b$.
- Indeed, from $f(1,1) = 1$ and $f(0,1) = 1$, due to scale invariance with $\lambda = b$, we conclude that

$$f(0,b) = b \text{ and } f(1,b) = b.$$

- Due to monotonicity, $0 \le a \le b$ implies that $b = f(0,b) \le f(a,b) \le f(b,b) = b$, thus $f(a,b) = b$.
- Due to commutativity, we now have $f(a,b) = \max(a,b)$ for all a and b.

18. Proof: Case When $f(1,1) \neq 1$

- Let us denote $v(k) \stackrel{\text{def}}{=} f(1, f(\ldots, 1)\ldots)$ (k times).
- Let us take $v(m \cdot n) = f(1, f(\ldots, 1)\ldots)$ ($m \cdot n$ times).
- We can divide the 1s into m groups with n 1s in each.
- Due to associativity,

  $$v(m \cdot n) = f(f(1, f(\ldots, 1)\ldots), \ldots, f(1, f(\ldots, 1)\ldots)).$$

- For each group, we have $f(1, f(\ldots, 1)\ldots) = v(n)$.
- Thus, $v(m \cdot n) = f(v(n), f(\ldots, v(n))\ldots)$ (m times).
- We know that $f(1, f(\ldots, 1)\ldots)$ (m times) $= v(m)$.
- Thus, by using scale-invariance with $\lambda = v(n)$, we conclude that $v(m \cdot n) = v(m) \cdot v(n)$.
- In particular, for every number p and n, we have

  $$v(p^n) = (v(p))^n.$$

19. Case $f(1,1) \neq 1$ (cont-d)

- If $v(2) = f(1,1) > 1$, then by monotonicity, we get $v(3) = f(1, v(2)) \geq f(1,1) = v(2)$.
- In general, we get $v(n+1) \geq v(n)$.
- In this case, the sequence $v(n)$ is (non-strictly) increasing.
- Similarly, if $v(2) = f(1,1) < 1$, then we get

$$v(3) \leq v(2).$$

- In general, we get $v(n+1) \leq v(n)$.
- In this case, the sequence $v(n)$ is strictly decreasing.
- Let us consider these two cases one by one.

20. $f(1,1) \neq 1$ and $v(n)$ Is Increasing

- Let us first consider the case when the sequence $v(n)$ is increasing.
- In this case, for every three integers m, n, and p, if $2^m \le p^n$, then $v(2^m) \le v(p^n)$, i.e., $(v(2))^m \le (v(p))^n$.
- For all m, n, and p, the inequality $2^m \le p^n$ is equivalent to $m \cdot \ln(2) \le n \cdot \ln(p)$, i.e., to $\dfrac{m}{n} \le \dfrac{\ln(p)}{\ln(2)}$.
- Similarly, the inequality $(v(2))^m \ge (v(p))^n$ is equivalent to $\dfrac{m}{n} \le \dfrac{\ln(v(p))}{\ln(v(2))}$.
- Thus, "if $2^m \le p^n$ then $(v(2))^m \le (v(p))^n$" takes the following form:

$$\text{for every rational } \frac{m}{n},\ \text{if } \frac{m}{n} \le \frac{\ln(p)}{\ln(2)} \text{ then } \frac{m}{n} \le \frac{\ln(v(p))}{\ln(v(2))}.$$

21. $f(1,1) \neq 1$ and $v(n)$ Is Increasing (cont-d)

- Similarly, for all m', n', and p, if $p^{n'} \le 2^{m'}$, then $v(p^{n'}) \le v(2^{m'})$, i.e., $(v(p))^{n'} \le (v(2))^{m'}$.
- The inequality $p^{n'} \le 2^{m'}$ is equivalent to $n' \cdot \ln(p) \le m' \cdot \ln(2)$, i.e., to $\dfrac{\ln(p)}{\ln(2)} \le \dfrac{m'}{n'}$.
- Also, the inequality $(v(p))^{n'} \le (v(2))^{m'}$ is equivalent to
$$\frac{\ln(v(p))}{\ln(v(2))} \le \frac{m'}{n'}.$$
- Thus, "if $p^{n'} \le 2^{m'}$ then $(v(p))^{n'} \le (v(2))^{m'}$" takes the following form:
for every rational $\dfrac{m'}{n'}$, if $\dfrac{\ln(p)}{\ln(2)} \le \dfrac{m'}{n'}$ then $\dfrac{\ln(v(p))}{\ln(v(2))} \le \dfrac{m'}{n'}$.
- Let us denote $\gamma \stackrel{\text{def}}{=} \dfrac{\ln(p)}{\ln(2)}$ and $\beta \stackrel{\text{def}}{=} \dfrac{\ln(v(p))}{\ln(v(2))}$.

22. $f(1,1) \neq 1$ and $v(n)$ Is Increasing (cont-d)

- For every $\varepsilon > 0$, there exist rational numbers $\dfrac{m}{n}$ and $\dfrac{m'}{n'}$ for which $\gamma - \varepsilon \le \dfrac{m}{n} \le \gamma \le \dfrac{m'}{n'} \le \gamma + \varepsilon$.
- The above two properties imply that $\dfrac{m}{n} \le \beta$ and $\beta \le \dfrac{m'}{n'}$ and thus, that $\gamma - \varepsilon \le \beta \le \gamma + \varepsilon$, i.e., that

$$|\gamma - \beta| \le \varepsilon.$$

- This is true for all $\varepsilon > 0$, so we conclude that $\beta = \gamma$, i.e., that $\dfrac{\ln(v(p))}{\ln(v(2))} = \gamma$.
- Hence, $\ln(v(p)) = \gamma \cdot \ln(p)$ and thus, $v(p) = p^\gamma$ for all integers p.

23. $f(1,1) \neq 1$: General Case

- We can reach:
  - a similar conclusion $v(p) = p^\gamma$ when the sequence $v(n)$ is decreasing and $v(2) < 1$, and
  - a conclusion that $v(p) = 0$ if $v(2) = 0$.
- By definition of $v(n)$, we have

$$f(v(m), v(m')) = v(m + m').$$

- Thus, we have $f(m^\gamma, (m')^\gamma) = (m + m')^\gamma$.
- By using scale-invariance with $\lambda = n^{-\gamma}$, we get

$$f\left(\frac{m^\gamma}{n^\gamma}, \frac{(m')^\gamma}{n^\gamma}\right) = \frac{(m + m')^\gamma}{n^\gamma}.$$

- Thus, for $a = \dfrac{m^\gamma}{n^\gamma}$ and $b = \dfrac{(m')^\gamma}{n^\gamma}$, we get $f(a,b) = (a^\alpha + b^\alpha)^{1/\alpha}$, where $\alpha \stackrel{\text{def}}{=} 1/\gamma$.

24. $f(1,1) \neq 1$: General Case (cont-d)

- Rational numbers $r = \dfrac{m}{n}$ are everywhere dense on the real line.
- Hence the values r^γ are also everywhere dense.
- So, every real number can be approximated, with any given accuracy, by such numbers.
- Thus, continuity implies that $f(a,b) = (a^\alpha + b^\alpha)^{1/\alpha}$ for every two real numbers a and b.
- The proposition is proven.