A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium

Alex Edmans, Xavier Gabaix and Augustin Landier

Wharton; NYU Stern and NBER; NYU Stern

Society for Economic Dynamics
July 2008

Introduction

- What should efficient CEO incentives look like? Our approach:

1. Introduce multiplicative preferences in P-A problem
   - This gives us optimal split of total pay into cash and shares
2. Market equilibrium
   - Endogenize total pay by embedding P-A problem into Gabaix and Landier (2008) competitive assignment model
     * Talent determines salary and size of one's firm
     * Model thus generates quantitative predictions for level of incentives and scaling with firm size

- End result: tractable, closed-form, neoclassical model of both incentives and pay under optimal contracting
- Multiplicative preferences allow us to match a number of stylized facts inconsistent with an additive model

Literature Review

- Calibrated models: Haubrich (1994), Garicano and Hubbard (2005), Gayle and Miller (2007), Dittmann and Maug (2007), Armstrong, Larcker and Su (2007), Coles, Lemmon and Meschke (2007)
- Baker and Hall (2004) inversion model to back out production function (preferences are always additive)
- Other market equilibrium models: Baranchuk, Macdonald and Yang (2007), Falato and Kadyrzhanova (2007)
  - This paper: multiplicative preferences and focus on size scalings
- Bebchuk and Fried (2004): pay reflects rent extraction, not efficiency. Kuhnen and Zwiebel (2007): theory and supportive empirics
  - But current practices pay be optimal: level of pay (Gabaix and Landier (2008)), severance pay (Almazan and Suarez (2003), Manso (2006), Inderst and Mueller (2006)), pensions (Edmans (2007)), perks (Rajan and Wulf (2006)

Basic Model: Incentive Pay in Partial Equilibrium

- Risk neutral, linear contract, binary effort

$$U = E\left[c\ g\left(e\right)\right]$$

$$P_1 = S\left(1+\eta\right)\left(1+e-\overline{e}\right)$$

- $c \geq 0$ is cash, $e \in \{0, \overline{e}\}$ is "effort"
  - $g\left(\overline{e}\right) = 1 < g\left(0\right) = 1/\left(1-\Lambda\overline{e}\right)$, where $0 < \Lambda\overline{e} < 1$.
- Firm chooses $c = f + \nu P_1$ where f is fixed salary, ν is # shares
  - Optimal contract elicits $e = \overline{e}$ and pays $E\left[c\right] = w$ (currently exogenous)
- **Proposition**: pay $\Lambda\%$ $(= \$\Lambda w)$ in shares, $(1-\Lambda)\%$ in cash
  - If wage doubles, $ shares must double
  - Alternative intuition: multiplicative model, so %/% incentives are relevant

Incentive Pay in Market Equilibrium

- Gabaix and Landier (2008) competitive assignment model
- $n \in [0, N]$ has size $S(n)$; manager $m \in [0, N]$ has talent $T(m)$. $S'(n) < 0,\ T'(m) < 0$
- Equilibrium involves $m = n$

$$w(n) = D(n_*)\, S(n_*)^{2/3} S(n)^{1/3}$$

- Optimal contract in market equilibrium:
  - Total pay: $w_n = D_* {S_*}^{1-\rho} {S_n}^{\rho}$
  - Stock: $\nu_n^* P_n = w_n \Lambda$
  - Cash: $f_n^* = w_n (1 - \Lambda)$

Pay-Performance Sensitivities in Market Equilibrium

	b^I	b^{II}	b^{III}
Measures	$\frac{\Delta \ln \text{Pay}}{\Delta \text{Firm Return}}$	$\frac{\Delta \$\text{Pay}}{\Delta \$\text{Firm Value}}$	$\frac{\Delta \$\text{Pay}}{\Delta \text{Firm Return}}$
Real vars	$\frac{\$\text{shares}}{\text{total pay}}$	% shares	$ shares
Used by	Murphy ('85)	Jen-Mur ('90), Sch ('98)	Holmstrom ('92)
	Rosen ('92)	Yermack ('95)	Hall-Lieb ('98)
Here	Λ	$\Lambda \frac{w}{S}$	Λw
$\propto S$	$b^I \propto S^0$	$b^{II} \propto S^{-2/3}$	$b^{III} \propto S^{1/3}$
$\propto S_*$	$b^I \propto S^0 S_*^0$	$b^{II} \propto S^{-2/3} S_*^{2/3}$	$b^{III} \propto S^{1/3} S_*^{2/3}$

- $\frac{\partial c / w}{\partial S / S} = \Lambda$ independent of size. Scale-independent contracts

Wealth-Performance Sensitivities

- Jensen and Murphy (1990), Hall and Liebman (1998), Core, Guay and Verrecchia (2003): portfolio incentives $\gg$ pay incentives
- Wealth-Performance Sensitivity analogs:

$$\begin{aligned} B^{I} &= \frac{\partial W}{\partial r}\frac{1}{w} = \frac{\Delta\$\text{wealth} \ / \ \Delta r}{\$\text{wage}} = \Lambda \frac{W}{w} \propto S^{0} \\ B^{II} &= \frac{\partial W}{\partial S} = \frac{\Delta\$\text{wealth}}{\Delta\$\text{Firm Value}} = \Lambda \frac{w}{S} \propto S^{-2/3} \\ B^{III} &= \frac{\partial W}{\partial r} = \frac{\Delta\$\text{wealth}}{\Delta r} = \Lambda W \propto S^{1/3} \end{aligned}$$

- Same scalings with S and S_* as before
- %/% incentives remain relevant

Application 1: Incentive Pay and Firm Size

- Model predicts $B^I \propto S^0$; $B^{II} \propto S^{-2/3}$; $B^{III} \propto S^{1/3}$

	$\ln(B^I)$	$\ln(B^{II})$	$\ln(B^{III})$
ln(Size)	0.038	-0.60	0.40
	(0.068)	(0.052)	(0.052)
N	6,470	6,470	6,470
$\overline{R}^2$	0.13	0.32	0.34

- Negative scaling of B^{II} with size (Jensen and Murphy (1990), Schaefer (1998)) is fully consistent with optimal contracting
- B^I [=(Δ wealth / Δ return) / wage] is size-independent
  - Desirable for comparisons between firms and over time
  - "Purer" measure of incentives for use as a control
- Multiplicative forms are necessary as well as sufficient to predict size-independent B^I

Application 2: Level of Incentive Pay

- We find $B^I \simeq 9$. In our paper, $B^I = \Lambda \frac{W}{w}$
- Utility from shirking (as fraction of wealth) is $\Lambda |\overline{e}| = B^I \frac{w}{W} |\overline{e}|$
- Plausible $|\overline{e}|$ is 10% (conservative given 30% takeover premium)
- Hence $\Lambda |\overline{e}| = 0.9w/W$, i.e. $\$0.9w$. Shirking can increase CEO's utility by no more than 0.9 times his pay

Application 3: Perks

- These are additive in firm value
- Can incentives prevent purchase of a $\overline{L} = \$10m$ jet?
  - Take market cap $S = \$10b$. Then, the jet lowers returns by $\overline{L}/S = 0.1\%$
- CEO's equity stake does not decline sufficiently to deter perks
  - Giving the CEO rents (in the form of extra shares) will not help
- Incentive should only be used to address large agency problems, not small ones (vs. Jensen and Meckling (1976))
  - Perks are best controlled through monitoring. Incremental role for active governance, consistent with Gompers et al. (2003), Yermack (2006)

Application 4: Effect of Risk on Incentives

- Classical models feature a trade-off between costs and benefits of incentives
  - Higher risk increases cost of incentives, so optimal level is lower
- Here, there is no trade-off. Under quite weak assumptions, maximum effort level is optimal. Hence
  - Incentives are independent of risk (Prendergast (2002))
  - Wealth volatility is increasing in risk: Table 3
- Absence of trade-off results from
  - Multiplicative production function
  - Maximum effort level

Detail-Independent Optimal Contracts (Edmans and Gabaix)

- In EGL, $B^I = \Lambda$ and is scale-independent. Extend to general contracts, RA, noise, continuous effort
- Continuous time: $dP_t / P_t = (r_f + \pi + a_t - \overline{a})\, dt + \sigma_t dZ_t$
- Let $R = P_1 / P_0$ be gross return. Optimal unrestricted contract is

$$c = kR^{\Lambda}$$

  independent of distribution of noise, utility function, reservation utility
- At $t = 0$, give CEO a portfolio of value $E[c]$, of which Λ is invested in stock and remainder in cash. Rebalanced continuously until $t = 1$
  - Combines tractability of Holmstrom and Milgrom (1987) with generality of Grossman and Hart (1983)
  - Consistent with simple contracts in reality

Discrete Time, General Case

- Assume noise occurs before action, as in Laffont and Tirole (1986) and Baker (1992)
- $r = a + \eta$
  - $a \in (\underline{a}, \overline{a}]$ is unobservable effort
  - η is noise (no restrictions except interval support).
- Utility function: $E\left[u\left(v\left(c\right) - g\left(a\right)\right)\right]$
  - $g\left(\cdot\right)$ is cost of effort
  - $u\left(\cdot\right)$ is a general utility function; $\underline{u}$ is reservation utility
  - $v\left(\cdot\right)$ is felicity function from cash ($\ln(c)$ in above example)
- Principal writes $\widetilde{c}(r, M)$ to implement a^*

Theorem

Unrestricted optimal contract:

$$c = v^{-1}\left(g'\left(a^*\right) r + K\right)$$

Extensions

- Contract form remains the same in a multiperiod model, and with multidimensional signals and actions
- Target action can depend on noise: $a = A(\eta)$
  - If firm is sufficiently large, $a^* = \bar{a}$ regardless of observed noise

Conclusion

- Neoclassical market equilibrium model endogenizing both total pay and its incentive component
- Modeling contributions:
  - Multiplicative functional forms
  - Embeds partial equilibrium incentive model into a competitive matching model of CEO talent to obtain a full market equilibrium
  - Tractable, closed form solutions (allowing empirical analysis)

Conclusion (cont'd)

- 4 applications:
  1. Understand incentives across firms. $B^{II} = (\Delta\text{Wealth}/\Delta\text{Firm value})$ should decline as $S^{-2/3}$.
     - * $B^{I} = (\Delta\text{Wealth}/\Delta r)$ / (pay) sensitivity is stable. Potentially attractive as an empirical measure
     - * Multiplicative functional forms necessary to match data
  2. Calibrate level of incentives. Level of WPS is fine if benefit from shirking $<$ annual pay
  3. Incentives are ineffective at solving perks
  4. Incentives are independent of risk
- Potential building block for more complex models. Simple market equilibrium that matches: (i) scale independence of contracts, (ii) scaling of B^{I}, B^{II}, B^{III}, (iii) positive relation between wealth volatility and σ, (iv) link between pay, size, aggregate size (GL).

Continuous Time - The General Case

- CT: $r_t = \int_0^t a_s ds + \eta_t$
  - $a \in (\underline{a}, \overline{a}]$ is unobservable effort
  - η is noise (no restrictions except interval support).
- Utility function: $E\left[u\left(v\left(c\right) - \int_0^T g\left(a_t\right) dt\right)\right]$
  - $g\left(\cdot\right)$ is cost of effort
  - $u\left(\cdot\right)$ is a general utility function; $\underline{u}$ is reservation utility
  - $v\left(\cdot\right)$ is felicity function from cash ($\ln(c)$ in above example)

Continuous Time - The General Case, Contract

Theorem

Unrestricted optimal contract:

$$c = v^{-1}\left(g'\left(a^*\right) r_T + K\right),$$

K satisfies $E\left[u\left(\cdot\right)\right] \geq \underline{u}$ with equality

- Allow a^* to Like EGL, closed-form optimal incentive level, constant % of output. Slope depends only on unit cost of effort, $\Lambda = g'\left(a^*\right) / A'(a^*)$. Details only affect K
  - Extends tractable contracts of Holmstrom and Milgrom (1987) to more general settings without exponential utility and Gaussian noise
  - Contrasts complexity of standard contracting models (e.g. Grossman and Hart (1983)); consistent with reality
- As in EGL, target effort level itself is DI

The Positive Relation between Wealth Volatility and Firm Volatility

	Ex ante measure of volatility			Ex post measure of volatility		
	$B^{I}\sigma_r$	$B^{II}\sigma_r$	$B^{III}\sigma_r$	$\frac{\lvert W_{t+1}-W_t\rvert}{w_t}$	$\frac{\lvert W_{t+1}-W_t\rvert}{S_t}$	$\lvert W_{t+1}-W_t\rvert$
σ_r	1.00	1.36	1.36	0.68	0.99	1.06
	(0.13)	(0.12)	(0.12)	(0.16)	(0.14)	(0.14)
S	0.02	-0.58	0.42	-0.05	-0.62	0.30
	(0.058)	(0.048)	(0.048)	(0.064)	(0.051)	(0.052)
N	6,276	6,276	6,276	4,545	4,545	4,545
$\overline{R}^2$	0.22	0.45	0.45	0.14	0.32	0.29

(All variables are in logs. Regressions are with year, industry fixed-effects)

Discrete Time, Multiple Periods

- Multiple discrete periods $1, ..., T$. Agent consumes c at time T
- In each period t:
  - n_t is realized
  - a_t is taken
  - $r_t = A_t(a_t) + n_t$ is observed
- Utility function $E\left[u\left(v\left(c\right) - \sum_{t=1}^{T} g\left(a_t\right)\right)\right]$
- Optimal contract: $c = v^{-1}\left(\sum_{t=1}^{T} \frac{g'(a_t^*)}{A_t'(a_t^*)} r_t + K\right)$

Discrete Time, One Period, CEO Incentives

- Contract can now be implemented with firm's securities
- $R = P_1 / P_0$ is gross return, $r = \ln R$ is log return
- $P_1 = V_0(1+a)(1+\varepsilon)$
- Specialize $v(c) = \ln c$, so utility function is $U\left(ce^{-g(a)}\right)$: multiplicative
- Optimal unrestricted contract is

$$c = kR^{\Lambda}$$

- At $t = 0$, give CEO a portfolio of value $E[c]$, of which Λ is invested in stock and remainder in cash. Rebalanced continuously until $t = 1$

Continuous Time, CEO Incentives

- $dP_t/P_t = (r_f + \pi + a_t - \overline{a})\, dt + \sigma_t dz_t$
- $E\left[U\left(c \exp\left(-\int_0^T g(a_t)\, dt\right)\right)\right]$
- $c = kR^{\Lambda}$ holds as before

The Discrete Time Analog

- Assume noise occurs before action, as in Laffont and Tirole (1986) and Baker (1992)
- Utility function: $E\left[u\left(v\left(c\right)-g\left(a\right)\right)\right]$
- Principal writes $c(r, M)$ to implement a^*

Theorem

Unrestricted optimal contract:

$$c = v^{-1}\left(\frac{g'\left(a^*\right)}{A'\left(a^*\right)}r + K\right)$$

- Hidden information (agent learns n before a) but no messages. Since a^* is implemented in all cases, one-to-one correspondence between r and n, so messages redundant

Empirical Implications

- Prendergast (2002): incentives are unrelated to risk
  - Indeed, we predict contracts are independent of risk and other details
- Dittmann and Maug (2007): observed contracts are convex, but a standard P-A model predicts concave contracts
  - Dittmann, Maug and Spalt (2008) predict convex contracts with a loss averse utility function
  - We can obtain convex contracts with a standard utility function

Further Extensions

- Contract is also robust to:
  - Endogenous target level a^*. Under quite weak assumptions, $a^* = \overline{a}$
  - Optimal action depends on noise
  - Agent can affect mean as well as volatility of payoffs
  - Multidimensional signals and actions

Conclusion: EG

- Contract still depends on Λ only even under general risk aversion and noise, and unrestricted contracting
  - Tractability is possible in more general settings than Holmstrom and Milgrom (1987)
  - Holds in both DT (if noise is before action) and CT
- Incentives continue to be independent of risk, and are also independent of other details
- Straightforward implementation for CEOs

Applications

1. Understand scaling of CEO incentives with firm size
   - Stylized fact that \$/\$ incentives fall with size (e.g. Jensen and Murphy (1990), Schaefer (1998))
     * Linear models: \$/\$ incentives should be constant across CEOs. Hence Bebchuk and Fried (2004): higher governance problems in larger firms
     * This paper: fully consistent with optimal contracting. Model predicts elasticity of $1/3 - 1 = -2/3$, we find -0.60
   - %/% incentives may be a desirable empirical measure
2. Calibrate level of incentives
   - Jensen and Murphy (1990): CEOs gain \$3.25 for every \$1,000 increase in shareholder wealth
   - Consistent with optimal contracting
3. Understand what issues can and cannot be solved with incentives
4. Understand "tenuous trade-off between risk and incentives" (Prendergast (2002))