# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Hope, Thomas M.H. and Friston, Karl J. and Price, Cathy J. and Leff, Alex P. and Rotshtein, Pia and Bowman, Howard (2018) Recovery After Stroke: Not So Proportional After All? Brain . ISSN 0006-8950. (In press)

### DOI

### Link to record in KAR

https://kar.kent.ac.uk/69746/

### Document Version

Author's Accepted Manuscript

# Recovery after stroke: not so proportional after all?

Thomas M.H. Hope [1] *, Karl Friston [1], Cathy J. Price [1], Alex P. Leff [2,3],

Pia Rotshtein [4] & Howard Bowman [4,5]

1. Wellcome Centre for Human Neuroimaging, University College London, UK.
2. Institute of Cognitive Neuroscience, University College London, UK.
3. Dept. of Brain Repair and Rehabilitation, Institute of Neurology, University College London, UK.
4. School of Psychology, University of Birmingham, UK.
5. School of Computing, University of Kent, UK.

* Corresponding author: Dr Thomas M.H. Hope, Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, WC1N 3AR, UK. Phone: +44 (0)20 3448 4345. E-mail: t.hope@ucl.ac.uk

**References:** 29

**Words:** (abstract): 196; (body): 3,156

**Keywords:** proportional recovery, stroke, methods, statistics.

**ABSTRACT**

The proportional recovery rule asserts that most stroke survivors recover a fixed proportion of lost function. To the extent that this is true, recovery from stroke can be predicted accurately from baseline measures of acute post-stroke impairment alone. Reports that baseline scores explain more than 80%, and sometimes more than 90%, of the variance in the patients' recoveries, are rapidly accumulating. Here, we show that these headline effect sizes are likely inflated.

The key effects in this literature are typically expressed as, or reducible to, correlation coefficients between baseline scores and recovery (outcome scores minus baseline scores). Using formal analyses and simulations, we show that these correlations will be extreme when outcomes are less variable than baselines, which they often will be in practice regardless of the real relationship between outcomes and baselines. We show that these effect sizes are likely to be over-optimistic in every empirical study that we found, which reported enough information for us to make the judgement, and argue that the same is likely to be true in other studies as well. The implication is that recovery after stroke may not be as proportional as recent studies suggest.

## 1. INTRODUCTION

Clinicians and researchers have long known stroke patients' initial symptom severity is related to their longer term outcomes (Jongbloed, 1986). Recent studies have suggested that this relationship is stronger than previously thought: that most patients recover a fixed proportion of lost function. Studies supporting this 'proportional recovery rule' are rapidly accumulating (Stinear, 2017): in five studies since 2015 (Byblow *et al.*, 2015; Feng *et al.*, 2015; Winters *et al.*, 2015; Buch *et al.*, 2016; Stinear *et al.*, 2017b), researchers used the Fugl-Meyer scale to assess patients' upper limb motor impairment within two weeks of stroke onset ('baselines'), and then again either three or six months post-stroke ('outcomes'). The results were consistent with earlier observations (Prabhakaran *et al.*, 2007; Zarahn *et al.*, 2011) that most patients recovered ~70% of lost function. Taken together, these studies report highly consistent recovery in over 500 patients, across different countries with different approaches to rehabilitation, regardless of the patients' ages at stroke onset, stroke type, sex, or therapy dose (Stinear, 2017). And there is increasing evidence that the rule also captures recovery from post-stroke impairments of lower limb function (Smith *et al.*, 2017), attention (Marchi *et al.*, 2017; Winters *et al.*, 2017), and language (Lazar *et al.*, 2010; Marchi *et al.*, 2017), and may even apply generally across cognitive domains (Ramsey *et al.*, 2017). Even rats appear to recover proportionally after stroke (Jeffers *et al.*, 2018).

Strikingly, many of these studies report that the baseline scores predict 80%-90%, or more, of the variance in empirical recovery. When predicting behavioural responses in humans, these effect sizes are unprecedented. Recently, Winters and colleagues (2015) reported that recovery predicted from baseline scores explained 94% of the variance in the empirical recovery of 146 stroke patients. Like many related reports (Stinear, 2017), this study also reported a group of (65) 'non-fitters', who did not make the predicted recovery. But if non-fitters can be distinguished at the acute stage, as this and other studies suggest (Stinear, 2017), the implication is that we can predict most patients' recovery near-perfectly, given baseline scores alone. Stroke researchers are used to thinking of recovery as a complex, multi-factorial process (Nelson *et al.*, 2016). If the proportional recovery rule is as powerful as it seems, post-stroke recovery is simpler and more consistent than previously thought.

In what follows, we argue that the empirical support for proportional recovery is weaker than it seems. These results are typically expressed as, or reducible to, correlations between baselines and recovery (outcomes minus baselines). These analyses pose well-known challenges, which have been discussed by statisticians for decades (Lord, 1956; Oldham, 1962; Cronbach and Furby, 1970; Hayes, 1988; Tu *et al.*, 2005). Much of this discussion is focused on problems induced by measurement noise, and measurement noise is also the focus of the only prior application of that discussion to the proportional recovery rule (Krakauer and Marshall, 2015). Here, we argue that empirical studies of proportional recovery after stroke are likely confounded entirely regardless of measurement noise.

Our argument is that: (a) correlations between baselines and recovery are spurious when they are stronger than correlations between baselines and outcomes; (b) this is likely when outcomes are less variable than baselines; which (c) will often happen in practice, whether or not recovery is proportional. This argument follows from a formal analysis of correlations between baselines and recovery, which we introduce in section 2 and illustrate with examples. We then employ that analysis to re-examining the empirical support for the proportional recovery rule in section 3.

## 2. THE RELATIONSHIPS BETWEEN BASELINES, OUTCOMES, AND RECOVERY

For the sake of brevity, we define 'baselines' = X, 'outcomes' = Y, and 'change' (recovery) = Δ: i.e. Y minus X. The 'correlation between baselines and outcomes' is r(X,Y), and the 'correlation between baselines and change' is r(X,Δ). Finally, we define the 'variability ratio' as the ratio of the standard deviation (σ) of Y to the standard deviation of X: $\sigma_Y/\sigma_X$.

X and Y are construed as lists of scores, with each entry being the performance of a single patient at the specified time point. We assume that higher scores imply better performance, so r(X,Δ) will be negative if recovery is proportional (to lost function). One can equally substitute 'lost function' (e.g. maximum score minus actual score), for 'baseline score', but while this makes r(X,Δ) positive if recovery is proportional, it is otherwise equivalent.

### 2.1. *Strong correlations imply the potential for accurate predictions*

Strong correlations between any two variables typically imply that we can use either variable to predict the other. Out-of-sample predictions should tend toward the least-squares line defined by the original (in-sample) correlation. Some empirical studies employ this logic to derive 'predicted recovery' (pΔ) from the least-squares line for r(X,Δ), reporting r(pΔ,Δ) instead of r(X,Δ) (Winters *et al.*, 2015; Marchi *et al.*, 2017). Since the magnitudes of r(X,Δ) and r(pΔ,Δ) are the same by definition (see proposition 8, Appendix A, and Figure 1), the preference for either expression over the other is arguably cosmetic.

Nevertheless, the correlation between predicted and empirical data is a common measure of predictive accuracy: the stronger the correlation, the better the predictions. Very strong correlations are unusual when predicting behavioural performance in humans – both because behaviour itself is complex, and because of measurement noise in behavioural assessment. Once $r(p\Delta,\Delta) > \sim 0.95$, for example (Winters *et al.*, 2015), this prognostic problem has seemingly been 'solved' more accurately than many might have thought possible.

### *2.2. r(X,Δ) is spurious when stronger than r(X,Y)*

Recovery is precisely the difference between baselines and outcomes. When $r(X,\Delta)$ is strong, implying that we can predict recovery accurately given baselines, it is tempting to assume that we can also predict outcomes equally accurately, by simply adding predicted recovery to baselines. More formally, the assumption is that $r(X+p\Delta,Y) \approx r(p\Delta,\Delta)$. This assumption is wrong.

In fact, $r(X+p\Delta,Y) \approx r(X,Y)$ (see appendix A, proposition 8, and Figure 1). When recovery is predicted from baselines, the correlation between 'baselines plus predicted recovery' and outcomes, is never stronger than the correlation between baselines and outcomes. When $r(X,\Delta)$ is stronger than $r(X,Y)$, $r(X,\Delta)$ is *spurious*, because it encourages an over-optimistic impression of how predictable outcomes are, given baselines.

### *2.3. The canonical example of spurious r(X,Δ)*

The canonical example of spurious $r(X,\Delta)$ is when X and Y are independent random variables with the same variance: $\sigma_Y/\sigma_X \approx 1$ and $r(X,Y) \approx 0$, but $r(X,\Delta) \approx -0.71$ (Oldham, 1962). This $r(X,\Delta)$ suggests that we can predict recovery relatively well, but we cannot use 'predicted recovery' to predict outcomes equally well (see Figure 1).

--Figure 1--

Krakauer and Marshall (2015) recently argued that this scenario has little relevance to (most) empirical studies of recovery after stroke. This is because: (a) spurious $r(X,\Delta)$ only emerge here when $r(X,Y)$ is weak; and (b) empirical $r(X,Y)$ are usually strong, because X and Y are dependent, repeated measurements from the same patients. If spurious $r(X,\Delta)$ only or mainly emerged when $\sigma_Y/\sigma_X \approx 1$ and

$r(X,Y) \approx 0$, they might indeed be irrelevant in practice. Unfortunately, spurious $r(X,\Delta)$ also emerge in another scenario, which is very common in studies of recovery after stroke.

### 2.4. *Spurious r(X,Δ) are likely when $\sigma_Y/\sigma_X$ is small*

For any X and Y, it can be shown that:

$$r(X,\Delta) = \frac{\sigma_Y \,.\, r(X,Y) - \sigma_X}{\sqrt{\sigma_Y^2 + \sigma_X^2 - 2\,.\,\sigma_X\,.\,\sigma_Y.\, r(X,Y)}} \quad \text{(Equation 1)}$$

A formal proof of Equation 1 is provided in Appendix A (proposition 4 and theorem 1; also see (Oldham, 1962)); its consequence is that $r(X,\Delta)$ is a function of $r(X,Y)$ and $\sigma_Y/\sigma_X$. To illustrate that function, we performed a series of simulations (see Appendix B) in which $r(X,Y)$ and $\sigma_Y/\sigma_X$ were varied independently. Figure 1 illustrates the results: a surface relating $r(X,\Delta)$ to $r(X,Y)$ and $\sigma_Y/\sigma_X$. Figure 2 illustrates example recovery data at six points of interest on that surface.

--Insert Figures 2 and 3 and Table 1--

Point A corresponds to the canonical example of spurious $r(X,\Delta)$, introduced in the last section: i.e., $\sigma_Y/\sigma_X \approx 1$ and $r(X,Y) \approx 0$, but $r(X,\Delta) \approx -0.71$ (see Figure 3a). At point B, $\sigma_Y/\sigma_X \approx 1$ and $r(X,Y)$ is strong, so recovery is approximately constant (Figure 3b) and $r(X,\Delta) \approx 0$, consistent with the view that strong $r(X,Y)$ curtail spurious $r(X,\Delta)$ (Krakauer and Marshall, 2015). However the situation is more complex when $\sigma_Y/\sigma_X$ is more skewed.

When $\sigma_Y/\sigma_X$ is large, Y contributes more variance to Δ, and $r(X, \Delta) \approx r(X,Y)$; this is Regime 1. Points C and D illustrate the convergence (Figure 3c-d). Data like this might suggest recovery proportional to *spared* function. By contrast, when $\sigma_Y/\sigma_X$ is small, X contributes more variance to Y-X, and $r(X,\Delta) \approx r(X,-X)$: i.e. -1 (see appendix A, theorem 2); this is Regime 2, where the confound emerges. Point E corresponds to data predicted by the proportional recovery rule: all patients recover exactly 70% of lost function (Figure 2e). Here, $\sigma_Y/\sigma_X$ is already small enough (0.3) to be dangerous: after randomly shuffling Y, $r(X,Y) \approx 0$, but $r(X,\Delta)$ is almost unaffected (Point F, and Figure 3f). Even if patients do recover proportionally, in other words, empirical data may enter territory, on the surface in Figure 2, where spurious $r(X,\Delta)$ are likely.

### 2.5. $\sigma_Y/\sigma_X$ will often be small, whether or not recovery is proportional

Proportional recovery implies small $\sigma_Y/\sigma_X$, but small $\sigma_Y/\sigma_X$ does not imply proportional recovery; for example, constant recovery with ceiling effects will produce the same effect. To illustrate this, we ran 1,000 simulations in which: (i) 1,000 baseline scores are drawn randomly with uniform probability from the range 0-65 (i.e. impaired on the 66-point Fugl-Meyer upper-extremity scale); (ii) outcome scores were calculated as the baseline scores plus half the scale's range (33); and (iii) outcome scores greater than 66 were set to 66 (i.e. a hard ceiling). Mean r(X,Y) and r(X,Δ) were calculated both before and after shuffling the outcomes data for each simulation. After shuffling, r(X,Y) ≈ 0 and r(X,Δ) = -0.88: ceiling effects make $\sigma_Y/\sigma_X$ small enough to encourage spurious r(X,Δ). And just as importantly, before shuffling, r(X,Y) = 0.89 and r(X,Δ) = -0.90: even when r(X,Δ) is *not* spurious (because r(X,Y) is similarly strong), we cannot conclude that recovery is really proportional.

## 3. RE-EXAMINING THE EMPIRICAL LITERATURE ON PROPORTIONAL RECOVERY

The relationships between r(X,Y), r(X,Δ) and $\sigma_Y/\sigma_X$, merit a re-examination of the empirical support for the proportional recovery rule. In the only study we found, which reports individuals' behavioural data, Zarahn and colleagues (2011) consider 30 patients' recoveries from hemiparesis after stroke. Across the whole sample, r(X,Y) = 0.80 and r(X,Δ) = -0.49; after removing 7 non-fitters: r(X,Y) = 0.75 and r(X,Δ) = -0.95. Removing the non-fitters increases the apparent predictability of recovery but reduces the predictability of outcomes (and reduces $\sigma_Y/\sigma_X$ from 0.88 to 0.36). Notably, the residuals for both correlations are identical (see Figure 4), and in fact this is always true (see Appendix A, proposition 10). r(X,Δ) has the same errors as r(X,Y), but a larger effect size: r(X,Δ) is over-optimistic.

--Insert Figure 3--

We can also use Equation 1 to reinterpret studies that do not report individual patient data. One example is the first study to report proportional recovery from aphasia after stroke (Lazar *et al.*, 2010). Here, r(X,Δ) ≈ -0.9 and $\sigma_Y/\sigma_X$ ≈ 0.48; Equation 1 implies that r(X,Y) was either ~0.78 or zero. Similarly, in the recent study of proportional recovery in rats (Jeffers *et al.*, 2018), $\sigma_Y/\sigma_X$ ≈ 0.8, and r(X,Δ) ≈ -0.71; Equation 1 implies that r(X,Y) was either much stronger (>0.95) or considerably weaker (~0.29) than r(X,Δ). In both cases, r(X,Δ) tells us less than expected about how predictable outcomes really were, given baseline scores.

Many recent studies report inter-quartile ranges (IQRs), rather than standard deviations, for the baselines and outcomes of patients deemed to recover proportionally. Accepting some room for error, we can also estimate $\sigma_Y/\sigma_X$ from those IQRs. In one case (Winters *et al.*, 2015), r(X,Δ) = -0.97 and $\sigma_Y/\sigma_X$ = 0.158, while in another (Veerbeek *et al.*, 2018), $\sigma_Y/\sigma_X$ = 0.438 and r(X,Δ) ≈ -0.88. In both cases, Equation 1 implies that r(X,Δ) would be at least that strong as that reported, regardless of r(X,Y): here again, the headline effect sizes do not tell us how predictable outcomes actually are, given baseline scores.

Many studies in this literature only relate baselines to recovery through multivariable models (Buch *et al.*, 2016; Marchi *et al.*, 2017; Winters *et al.*, 2017); in these studies, we cannot demonstrate confounds directly with Equation 1. Nevertheless, these studies are also probably confounded, because any inflation in one variable's effect size will inflate the multivariable model's effect size as well. As discussed in section 2.5, empirical studies of recovery after stroke should tend to encourage small $\sigma_Y/\sigma_X$, whether or not recovery is really proportional. Consequently, the null hypothesis will rarely be that r(X,Δ) ≈ 0. For example, in the only multivariable modelling study, which reports IQRs for its fitter-patients' baselines and outcomes (Stinear *et al.*, 2017c), $\sigma_Y/\sigma_X$ ≈ 0.48, which implies that the weakest r(X,Δ) was -0.88, for any positive value of r(X,Y).

Finally, while r(X,Δ) can be misleading if it is extreme relative to r(X,Y), the reverse is also true. One study in this literature which employs outcomes as the dependent variable, rather than recovery (Feng *et al.*, 2015), reports that r(X,Y) ≈ 0.8 and $\sigma_Y/\sigma_X$ = 1.2 in their 'combined' group of 76 patients. By Equation 1, r(X,Δ) = -0.05: i.e. recovery was uncorrelated with baseline scores. These authors only report proportional recovery in a sub-sample of their patients (but not the information we need to re-examine that claim), but their full sample seems better described by constant recovery (as in Figure 3b).

## 4. Discussion

The proportional recovery rule is striking because it implies that recovery is simple and consistent across patients (non-fitters notwithstanding), and because that implication appears to be justified by strong empirical results (Stinear, 2017). We contend that the empirical support for the rule is weaker than it seems.

In summary, our argument is that r(X,Δ) is spurious when stronger than r(X,Y), and that the conditions which encourage spurious r(X,Δ) will be common in empirical studies of recovery after stroke, whether or not recovery is really proportional. Many empirical r(X, Δ) in this literature appear to be spurious in this sense. And in any case, strong r(X,Δ) are insufficient evidence for proportional recovery if they are *not* spurious (because they are accompanied by similarly strong r(X,Y)).

The only previous discussion of the risk of spurious r(X,Δ), in analyses of recovery after stroke, (Krakauer and Marshall, 2015), concluded that this risk is small provided the tools used to measure post-stroke impairment are reliable: i.e. so long as measurement noise is minimal. Crucially, our analysis applies entirely regardless of measurement noise. We contend that the risk of spurious r(X,Δ) is significant, if there are ceiling effects on the scale used to measure post-stroke impairment, and if most patients improve between baseline and subsequent assessments. The criteria will usually be met in practice, because every practical measurement of post-stroke impairment employs a finite scale, and because non-fitters, who do not make the predicted recovery, are removed prior to calculating r(X,Δ).

We are not suggesting that there is anything wrong with the practice of distinguishing fitters from non-fitters. Indeed, our results prove that this work may be valid regardless of our other concerns. Non-fitters do not recover as predicted; by definition, they contribute the largest, negative residuals to r(X,Δ). In Figure 4 and appendix A (proposition 9), we show that the residuals for r(X,Y) and r(X,Δ) are exactly the same, so the same patients will be placed in the same sub-groups regardless of which correlation is used, and biomarkers which distinguish those sub-groups at the acute stage (Stinear, 2017), will be equally accurate regardless of which correlation is used. Nevertheless, extreme r(X,Δ) for patients classified as fitters, will naturally encourage the assumption that those fitters' outcomes are largely determined by initial symptom severity. If this assumption is true, therapeutic interventions must be largely ineffective (or at least redundant) for these patients. Our analysis suggests that this assumption is wrong.

Nevertheless, we are not claiming that the proportional recovery rule is wrong. Our analysis suggests that empirical studies to date do not demonstrate that the rule holds, or how well, but we could only confirm that r(X,Δ) was actually over-optimistic in one study, which reported individual patient data. And while we have also shown that extreme r(X,Δ) and r(X,Y) can result from non-proportional (constant) recovery, this is simply a plausible alternative hypothesis about how patients really recover.

Quite how to interpret empirical recovery with confidence in this domain, remains an open question: we have articulated a problem here, hoping that recognition of the problem will motivate work to solve it. Nevertheless, we can make some recommendations for future studies in the field.

First, these studies should report r(X,Δ), r(X,Y), and $\sigma_Y/\sigma_X$, for those patients deemed to recover proportionally. Despite our concerns about r(X,Δ), we do learn something when r(X,Y) is strong, but r(X,Δ) is weak, as in Feng and colleagues' (2015) results in section 3, which appeared to be better explained by constant recovery than by proportional recovery.

Second, future studies should consider explicitly testing the hypothesis that recovery depends on baseline scores (Oldham, 1962; Hayes, 1988; Tu *et al.*, 2005; Tu and Gilthorpe, 2007; Chiolero *et al.*, 2013). These tests sensibly acknowledge that the null hypothesis is rarely r(X,Δ) ≈ 0 in these analyses. However, they do not address the proper measurement and interpretation of effect sizes, which is our primary concern here; somewhat paradoxically, this means that they may be less useful in larger samples than in smaller samples (Friston, 2012; Lorca-Puls *et al.*, 2018).

These hypothesis tests will also all be confounded by ceiling effects. We recommend that future studies should measure the impact of such effects, perhaps by reporting the shapes of the distributions of X and Y (greater asymmetry implying more prominent ceiling effects). Future studies should also attempt to minimise ceiling effects. One approach might be to remove patients whose outcomes are at ceiling: though certainly inefficient, this does at least remove the spurious r(X,Δ) in our simulations of constant recovery (section 2.5). However, it may be difficult to determine which patients to remove in practice; the Fugl-Meyer scale, for example, imposes item-level ceiling effects, which could distort $\sigma_Y/\sigma_X$ well below the maximum score. A better, though also more complex alternative, may be to employ assessment tools expressly designed to minimise ceiling effects, or to add such tools to those currently in use.

More generally, we may need to replace correlations with alternative methods, which can provide less ambiguous evidence for the proportional recovery rule. One principled alternative might employ Bayesian model comparison to adjudicate between different forward or generative models of the data at hand: i.e. using the empirical data to quantify evidence for or against competing hypotheses about the nature of recovery, which may or may not be conserved across patients. We hope that our analysis here will encourage work to develop such methods, delivering better evidence for (or against) the proportional recovery rule.

**ACKNOWLEDGMENTS**

This study was supported by the Medical Research Council (MR/M023672/1, MR/K022563/1), Wellcome (091593/Z/10/Z), and the Stroke Association (TSA PDF 2017/02). The funders had no participation in the design and results of this study. We would also like to thank the (anonymous) reviewers, whose constructive comments helped us to improve the paper.

**REFERENCES**

Buch ER, Rizk S, Nicolo P, Cohen LG, Schnider A, Guggisberg AG. Predicting motor improvement after stroke with clinical assessment and diffusion tensor imaging. Neurology 2016; 86(20): 1924-5.

Byblow WD, Stinear CM, Barber PA, Petoe MA, Ackerley SJ. Proportional recovery after stroke depends on corticomotor integrity. Annals of neurology 2015; 78(6): 848-59.

Chiolero A, Paradis G, Rich B, Hanley JA. Assessing the Relationship between the Baseline Value of a Continuous Variable and Subsequent Change Over Time. Frontiers in Public Health 2013; 1: 29.

Cronbach LJ, Furby L. How we should measure" change": Or should we? Psychological bulletin 1970; 74(1): 68.

Feng W, Wang J, Chhatbar PY, Doughty C, Landsittel D, Lioutas V-A*, et al.* Corticospinal tract lesion load: An imaging biomarker for stroke motor outcomes. Annals of neurology 2015; 78(6): 860-70.

Friston K. Ten ironic rules for non-statistical reviewers. NeuroImage 2012; 61(4): 1300-10.

Hayes RJ. Methods for assessing whether change depends on initial value. Statistics in medicine 1988; 7(9): 915-27.

Jeffers MS, Karthikeyan S, Corbett D. Does Stroke Rehabilitation Really Matter? Part A: Proportional Stroke Recovery in the Rat. Neurorehabilitation and neural repair 2018; 32(1): 3-6.

Jongbloed L. Prediction of function after stroke: a critical review. Stroke; a journal of cerebral circulation 1986; 17(4): 765-76.

Krakauer JW, Marshall RS. The proportional recovery rule for stroke revisited. Annals of neurology 2015; 78(6): 845-7.

Lazar RM, Minzer B, Antoniello D, Festa JR, Krakauer JW, Marshall RS. Improvement in aphasia scores after stroke is well predicted by initial severity. Stroke; a journal of cerebral circulation 2010; 41(7): 1485-8.

Lorca-Puls DL, Gajardo-Vidal A, White J, Seghier ML, Leff AP, Green DW*, et al.* The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. Neuropsychologia 2018; 115: 101-11.

Lord FM. THE MEASUREMENT OF GROWTH. ETS Research Bulletin Series 1956; 1956(1): i-22.

Marchi NA, Ptak R, Di Pietro M, Schnider A, Guggisberg AG. Principles of proportional recovery after stroke generalize to neglect and aphasia. European Journal of Neurology 2017; 24(8): 1084-7.

Nelson MLA, Hanna E, Hall S, Calvert M. What makes stroke rehabilitation patients complex? Clinician perspectives and the role of discharge pressure. Journal of Comorbidity 2016; 6(2): 35-41.

Oldham PD. A note on the analysis of repeated measurements of the same subjects. Journal of Chronic Diseases 1962; 15(10): 969-77.

Prabhakaran S, Zarahn E, Riley C, Speizer A, Chong JY, Lazar RM*, et al.* Inter-individual variability in the capacity for motor recovery after ischemic stroke. Neurorehabilitation and neural repair 2007; 22.

Ramsey LE, Siegel JS, Lang CE, Strube M, Shulman GL, Corbetta M. Behavioural clusters and predictors of performance during recovery from stroke. Nature human behaviour 2017; 1: 0038.

Smith M-C, Byblow WD, Barber PA, Stinear CM. Proportional Recovery From Lower Limb Motor Impairment After Stroke. Stroke; a journal of cerebral circulation 2017.

Stinear CM. Prediction of motor recovery after stroke: advances in biomarkers. The Lancet Neurology 2017; 16(10): 826-36.

Stinear CM, Barber PA, Petoe M, Anwar S, Byblow WD. The PREP algorithm predicts potential for upper limb recovery after stroke. Brain A J Neurol 2012; 135.

Stinear CM, Byblow WD, Ackerley SJ, Smith M-C, Borges VM, Barber PA. PREP2: A biomarker-based algorithm for predicting upper limb function after stroke. Annals of Clinical and Translational Neurology 2017a; 4(11): 811-20.

Stinear CM, Byblow WD, Ackerley SJ, Smith M-C, Borges VM, Barber PA. Proportional Motor Recovery After Stroke. Implications for Trial Design 2017b; 48(3): 795-8.

Stinear CM, Byblow WD, Ackerley SJ, Smith MC, Borges VM, Barber PA. Proportional Motor Recovery After Stroke: Implications for Trial Design. Stroke; a journal of cerebral circulation 2017c; 48(3): 795-8.

Tu YK, Baelum V, Gilthorpe MS. The relationship between baseline value and its change: problems in categorization and the proposal of a new method. European journal of oral sciences 2005; 113(4): 279-88.

Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: a review and evaluation. Statistics in medicine 2007; 26(2): 443-57.

Veerbeek JM, Winters C, van Wegen EEH, Kwakkel G. Is the proportional recovery rule applicable to the lower limb after a first-ever ischemic stroke? PLOS ONE 2018; 13(1): e0189279.

Warrington E, McKenna P. The graded naming test. London: NFER 1983.

Winters C, Wegen EEH, Daffertshofer A, Kwakkel G. Generalizability of the Proportional Recovery Model for the Upper Extremity After an Ischemic Stroke. Neurorehabilitation and neural repair 2015; 29.

Winters C, Wegen EEHv, Daffertshofer A, Kwakkel G. Generalizability of the Maximum Proportional Recovery Rule to Visuospatial Neglect Early Poststroke. Neurorehabilitation and neural repair 2017; 31(4): 334-42.

Zarahn E, Alon L, Ryan SL, Lazar RM, Vry M-S, Weiller C*, et al.* Prediction of motor recovery using initial impairment and fMRI 48 h poststroke. Cereb Cortex (New York, NY 1991) 2011; 21.

Figure 1

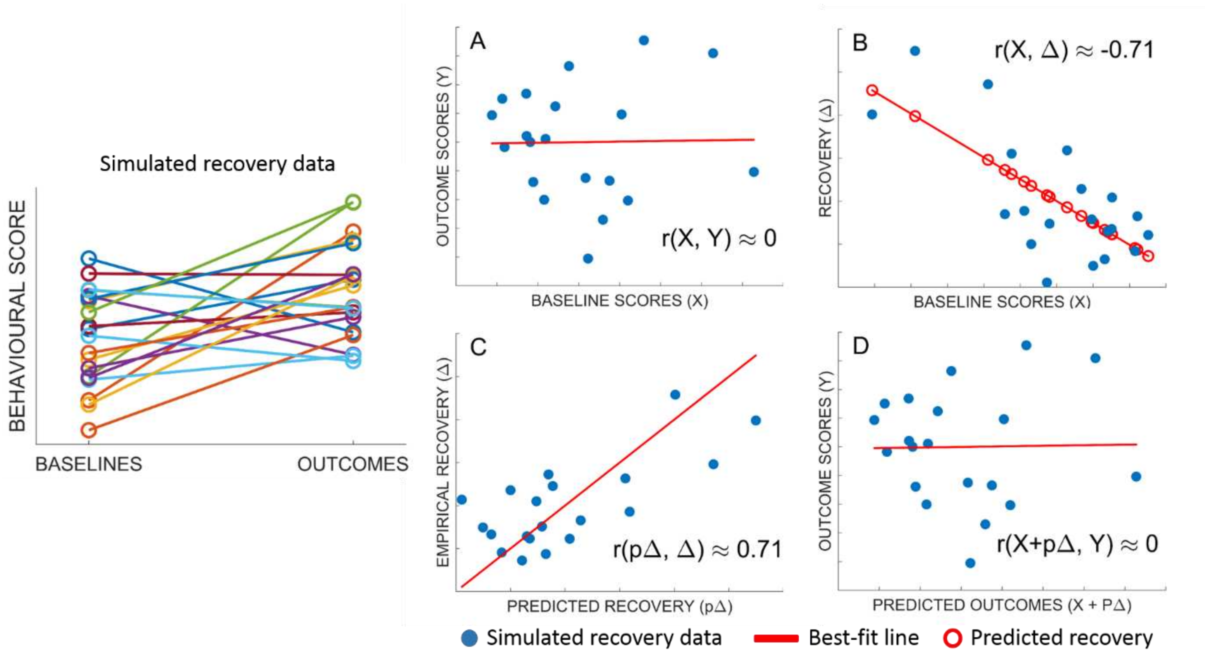


**Figure 1: A canonical example of spurious r(X, Δ).** Baselines scores are uncorrelated with outcomes (A), but baseline scores appear to be strongly correlated with recovery (B). That correlation can be used to derive predicted recovery, which is strongly correlated with empirical recovery (C) – but predicted outcomes, derived from that predicted recovery, are still uncorrelated with empirical outcomes (D).

Figure 2

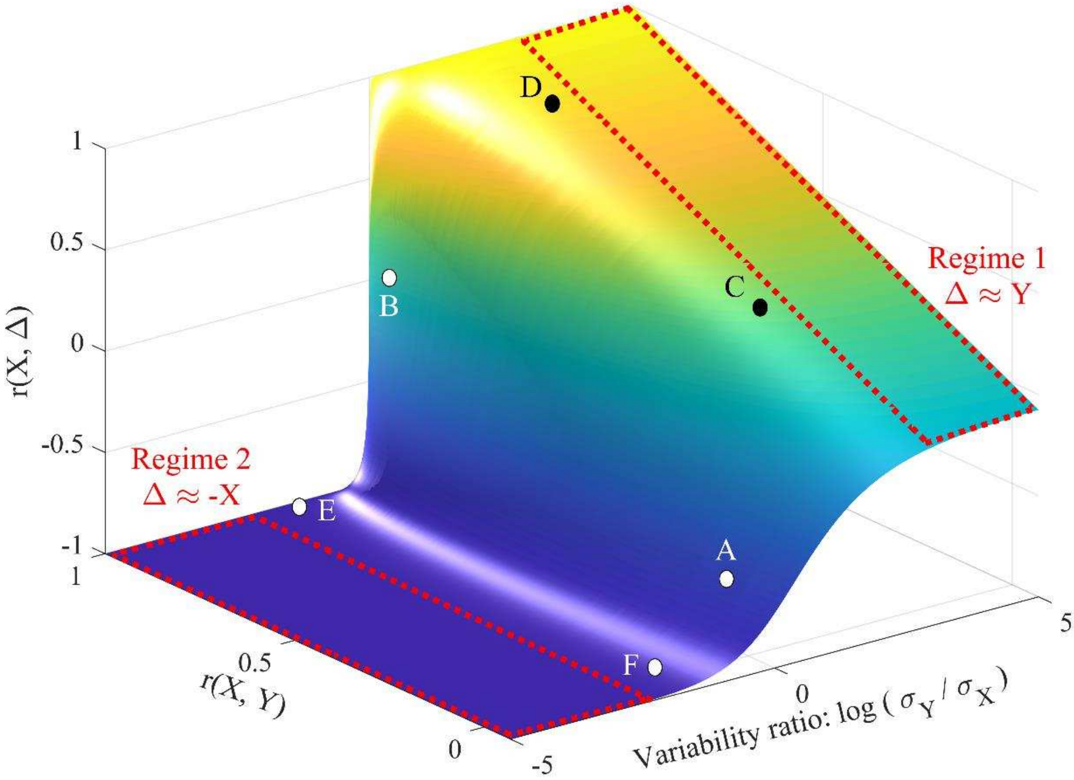


**Figure 2: The relationship between r(X,Y), r(X,Δ) and $\sigma_Y/\sigma_X$.** Note that the x-axis is log-transformed to ensure symmetry around 1; when X and Y are equally variable, $\log(\sigma_Y/\sigma_X) = 0$. Proposition 7 in Appendix A provides a justification for unambiguously using a ratio of standard deviations in this figure, rather than $\sigma_Y$ and $\sigma_X$ as separate axes. The two major regimes of Equation 1 are also marked in red. In Regime 1, Y is more variable than X, so contributes more variance to Δ, and r(X,Δ) ≈ r(X,Y). In Regime 2, X is more variable than Y, so X contributes more variance to Δ, and r(X,Δ) ≈ r(X,-X) (i.e. -1). The transition between the two regimes, when the variability ratio is not dramatically skewed either way, also allows for spurious r(X,Δ). For the purposes of illustration, the figure also highlights 6 points of interest on the surface, marked A-F; examples of simulated recovery data corresponding to these points are provided in Figure 3.

Figure 3

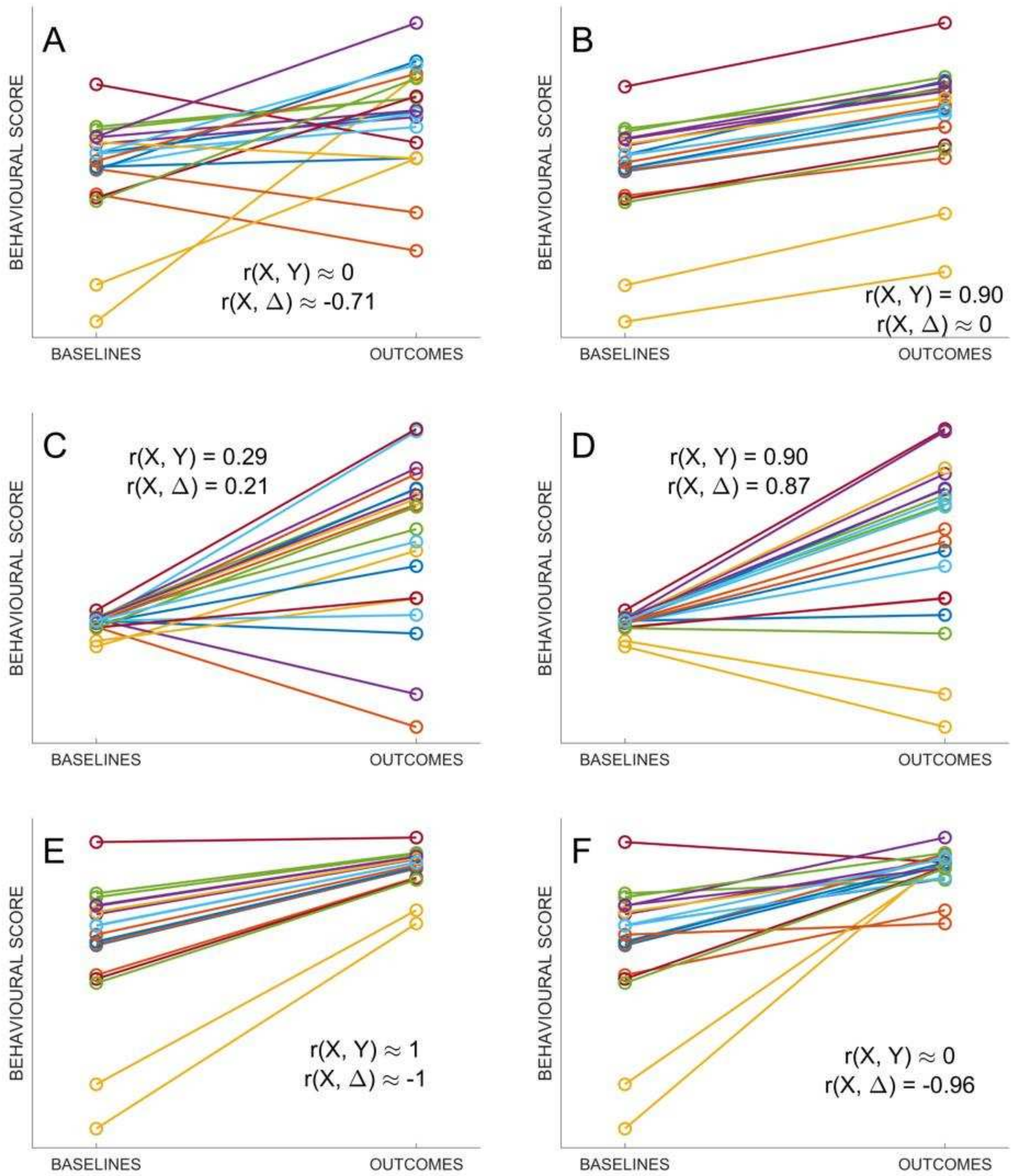


**Figure 3: Exemplar points on the surface in Figure 2.** Simulated recovery data, corresponding to the points A-F marked on the surface in Figure 1. (A) Baselines and outcomes are entirely independent ($r(X,Y)=0$), yet $r(X, \Delta)$ is relatively strong; this is the canonical example of mathematical coupling, first introduced by Oldham (1962); (B) Recovery is constant with minimal noise, so baselines and outcomes are equally variable ($\sigma_Y/\sigma_X \approx 1$) and recovery is unrelated to baseline scores ($r(X, \Delta) \approx 0$); (C-D) Outcomes are more variable than baselines ($\sigma_Y/\sigma_X \approx 5$), and $r(X, \Delta)$ converges to $r(X, Y)$; (E) Recovery is 70% of lost function, so outcomes are less variable than baselines ($\sigma_Y/\sigma_X \approx 0.3$); even with shuffled outcomes data (F) baselines and recovery still appear to be strongly correlated.

Figure 4

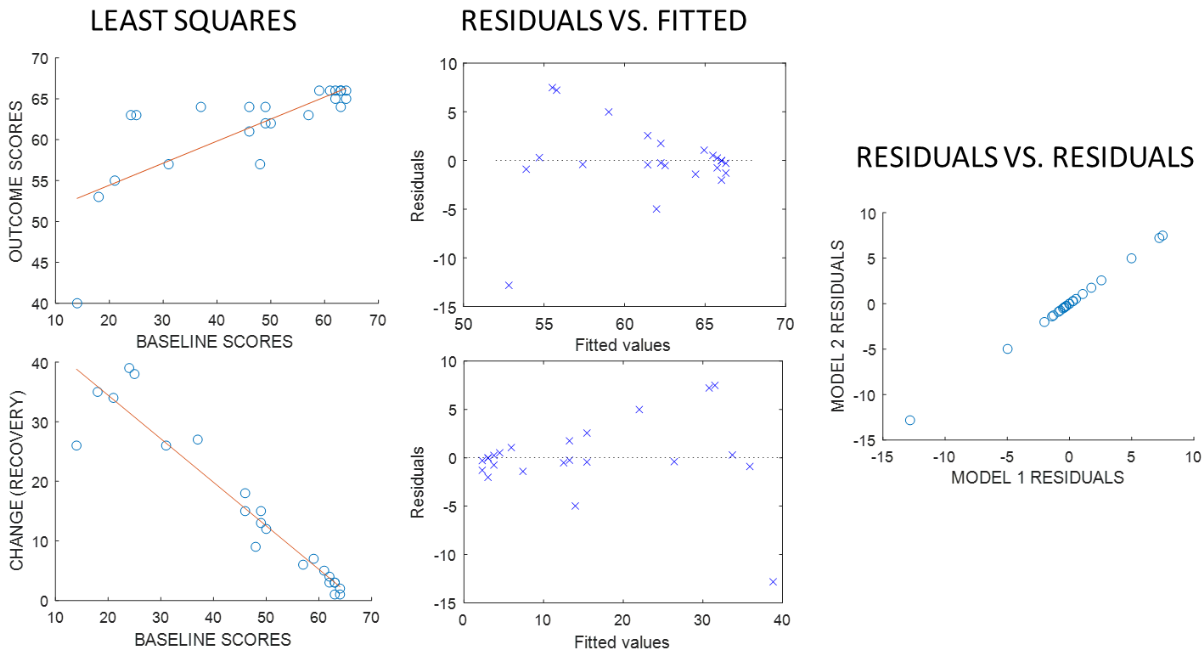


**Figure 4:** (Left) Least squares linear fits for analyses relating baselines to (upper) outcomes and (lower) recovery, using the fitters' data reported by Zarahn and colleagues (Zarahn *et al.*, 2011). (Middle) Plots of residuals relative to each least squares line, against the fitted values in each case. (Right) A scatter plot of the residuals from the model relating baselines to change, against the residuals from the model relating baselines to outcomes: the two sets of residuals are the same.

Table 1

| REGIME | VARIABILITY OF Y ($\sigma_Y$) | VARIABILITY OF X ($\sigma_X$) | Δ [= Y-X] | r(X,Δ) [=r(X,Y-X)] |
|---|---|---|---|---|
| 1 | Smaller | Larger | Y-X ≈ -X | r(X,Y-X) ≈ r(X,-X) = -1 |
| 2 | Larger | Smaller | Y-X ≈ Y | r(X,Y-X) ≈ r(X,Y) |

**Supplementary Appendix A: formal relationships between the correlations**

We present a simple, general and self-contained formulation of the proportional recovery concept. We have derived all of the key results from first principles, while acknowledging previous presentations of these results when they can be found in the literature.

We assume two variables $X'$ and $Y'$ corresponding to performance at initial test ($X'$) and at second test ($Y'$). These will be represented as column vectors, with each entry being the performance of a single patient and vector lengths being $N \in \mathbb{N}$. Performance improves as numbers get bigger, up to a maximum, denoted $Max$, which corresponds to no discernible deficit. Severity is measured as difference from maximum, i.e. $Max - X'$.

The two variables ($X'$ and $Y'$) could be specialised to more detailed formulations: e.g., true score theory or with an explicit modelling of measurement or state error. However, this would not impact any of the derivations or inferences that follow. Indeed, the results that we present would hold even in the complete absence of measurement noise, which has been considered the main concern for the validity of quantifications of proportional recovery.

## Demeaning

Without loss of generality, we work with demeaned variables. That is, where over-lining denotes mean, we define new variables as,

$$X = X' - \bar{X}'$$

$$Y = Y' - \bar{Y}'$$

This also means that recovery, i.e. $Y - X$, will be demeaned, since, using proposition 1, the following holds.

$$Y - X = (Y' - \bar{Y}') - (X' - \overline{X'}) = (Y' - X') - (\bar{Y}' - \overline{X'}) = (Y' - X') - \overline{(Y' - X')}$$

**Proposition 1**

Let $V$ and $W$ be vectors of the same length, denoted $N$. Then, the following holds,

$$\bar{V} + \bar{W} = \overline{(V + W)}$$

with $\bar{V} - \bar{W} = \overline{(V - W)}$ as a trivial consequence.

**Proof**

By distributivity of multiplication through addition and associativity of addition, the following holds.

$$\bar{V} + \bar{W} = \left(\frac{1}{N}\sum_{i=1}^{N} V_i\right) + \left(\frac{1}{N}\sum_{i=1}^{N} W_i\right) = \frac{1}{N}\left(\sum_{i=1}^{N} V_i + \sum_{i=1}^{N} W_i\right) = \frac{1}{N}\left(\sum_{i=1}^{N}(V_i + W_i)\right) = \overline{(V + W)}$$

QED

## Correlations

There are two basic correlations we are interested in, (1) the correlation between initial performance and performance at second test, i.e. $r(X, Y)$, and (2) the correlation between initial performance and recovery, i.e. $r(X, Y - X) = r(X, \Delta)$. The latter of these is the key relationship, and we would expect this to be a negative correlation; that is, as initial performance is smaller (i.e. further from $Max$), the larger is recovery. (One could also formulate the correlation as $r((Max - X), Y - X)$, which would flip the correlation to positive, but the two approaches are equivalent).

Our main correlations are defined as follows,

$$r(X, Y) = \frac{\sum_{i=1}^{N} X_i \, . \, Y_i}{(N-1) \, . \, \sigma_X \, . \, \sigma_Y}$$

$$r(X, (Y - X)) = \frac{\sum_{i=1}^{N} (X_i \, . \, (Y_i - X_i))}{(N-1) \, . \, \sigma_X \, . \, \sigma_{(Y-X)}}$$

## Standard Deviation of a Difference

We need a straightforward result on the standard deviation of a difference.

**Proposition 2**

$$\sigma_{(A-B)} = \sqrt{\sigma_A^2 + \sigma_B^2 - 2 \, . \, cov(A, B)}$$

**Proof**

The result is a direct consequence of the following standard result from probability theory, e.g. see Ross, S. M. (2014). *Introduction to probability and statistics for engineers and scientists*. Academic Press.,

$$\sigma_{(A-B)}^2 = \sigma_A^2 + \sigma_B^2 - 2 \, . \, cov(A, B)$$

## Key Results

The following proposition enables us to express the key correlation, $r(X, (Y - X))$, in terms of covariance of its constituent variables.

**Proposition 3**

$$r(X, (Y - X)) = \frac{cov(X, Y) - cov(X, X)}{\sigma_X \, . \sqrt{\sigma_Y^2 + \sigma_X^2 - 2 \, . \, cov(X, Y)}}$$

**Proof**

Using distributivity of multiplication through addition, associativity of addition, the definition of covariance and proposition 2, we can reason as follows.

$$r(X, (Y - X)) = \frac{\sum_{i=1}^{N} (X_i \, . \, (Y_i - X_i))}{(N-1) \, . \, \sigma_X \, . \, \sigma_{(Y-X)}} = \frac{\sum_{i=1}^{N} (X_i \, Y_i - \, X_i \, X_i)}{(N-1) \, . \, \sigma_X \, . \, \sigma_{(Y-X)}} = \frac{\sum_{i=1}^{N} (X_i \, Y_i) - \sum_{i=1}^{N} (X_i \, X_i)}{(N-1) \, . \, \sigma_X \, . \, \sigma_{(Y-X)}}$$

$$= \frac{cov(X, Y) - cov(X, X)}{\sigma_X \cdot \sigma_{(Y-X)}} = \frac{cov(X, Y) - cov(X, X)}{\sigma_X \, . \sqrt{\sigma_Y^2 + \sigma_X^2 - 2 \, . \, cov(X, Y)}}$$

QED

It is straightforward to adapt proposition 3 to be fully in terms of correlations.

**Proposition 4**

$$r(X,(Y-X)) = \frac{\sigma_Y\,.\,r(X,Y) - \sigma_X\,.\,r(X,X)}{\sqrt{\sigma_Y^2 + \sigma_X^2 - 2\,.\,\sigma_X\,.\,\sigma_Y.\,r(X,Y)}}$$

**Proof**

Straightforward from proposition 3 and definition of correlations, which gives the relationship $cov(A,B) = \sigma_A\,.\,\sigma_B\,.\,r(A,B)$. QED

## Scale Invariance

The next set of propositions justifies working with a standardised $X$ variable.

**Lemma 1**

$$\forall c \in \mathbb{R}\ \cdot |c|.\,\sigma_A = \sigma_{(c.A)}$$

**Proof**

Using distributivity of a multiplicative constant through averaging, $\sqrt{d^2} = |d|$ and distributivity of square root through multiplication, we can reason as follows.

$$\sigma_{(c.A)} = \sqrt{\frac{\sum_{i=1}^{N}\ (c.A_i - \overline{c.A})^2}{N-1}} = \sqrt{\frac{\sum_{i=1}^{N}\ (c.A_i - c.\bar{A})^2}{N-1}}\ =\ |c|\,.\sqrt{\frac{\sum_{i=1}^{N}\ (A_i - \bar{A})^2}{N-1}} = |c|.\,\sigma_A$$

QED

**Proposition 5 (Invariance to scaling)**

The absolute magnitude of a correlation is not changed by scaling either variable by a constant, i.e.

$$\forall c \in \mathbb{R}\ \cdot r(A,B) =\ sign(c).\,r(c.A,B) =\ sign(c).\,r(A,c.B)$$

where $sign(d)\ =\ if\ \ (d<0)\ \ then\ \ -1\ \ else\ \ +1$.

**Proof**

For any $c \in \mathbb{R}$, using distributivity of multiplication through mean and addition, and lemma 1, the following holds,

$$r(c.A,B) = \frac{\sum_{i=1}^{N}\ (c.A_i - \overline{c.A})(B_i - \bar{B})}{(N-1)\,.\,\sigma_{(c.A)}\,\sigma_B} = \frac{\sum_{i=1}^{N}\ (c.A_i - c.\bar{A})(B_i - \bar{B})}{(N-1)\,.\,\sigma_{(c.A)}\,\sigma_B}$$

$$= \frac{c\,.\sum_{i=1}^{N}\ (A_i - \bar{A})(B_i - \bar{B})}{(N-1)\,.\,|c|\,.\,\sigma_A\,.\,\sigma_B} = \frac{sign(c).\sum_{i=1}^{N}\ (A_i - \bar{A})(B_i - \bar{B})}{(N-1)\,.\,\sigma_A\,.\,\sigma_B} = sign(c).\,r(A,B)$$

Then, one can multiply both sides by $sign(c)$ to obtain $r(A,B) =\ sign(c).\,r(c.A,B)$. Additionally, as correlations are symmetric, $sign(c).\,r(c.B,A) =\ sign(c).\,r(A,c.B)$, and the full result follows.

QED

**Corollary 1**

$$\forall c \in \mathbb{R}\ \cdot r(A,B) =\ r(c.A,c.B)$$

**Proof**

Follows from twice applying proposition 5, and that $sign(c)^2 = +1$. QED

**Proposition 6**

$$\forall c \in \mathbb{R} \cdot r(X,(Y-X)) = r(c.X,(c.Y-c.X))$$

**Proof**

We can use distributivity of multiplication through subtraction and corollary 1 to give us the following.

$$r\big(c.X,(c.Y-c.X)\big) = r\big(c.X, c.(Y-X)\big) = r(X,(Y-X))$$

QED

It follows from proposition 6 that we can work with a standardised $X$ variable, since,

$$r(X/\sigma_X,(Y/\sigma_X - X/\sigma_X)) = r(X,(Y-X))$$

**Proposition 7 (Sufficiency of variability ratio)**

Assume two pairs of variables: $X_1, Y_1$ and $X_2, Y_2$, such that, $r(X_1,Y_1) = r(X_2,Y_2)$, then,

$$\frac{\sigma_{Y_1}}{\sigma_{X_1}} = \frac{\sigma_{Y_2}}{\sigma_{X_2}} \implies r\big(X_1,(Y_1-X_1)\big) = r\big(X_2,(Y_2-X_2)\big)$$

**Proof**

The proof has two parts.

1) We consider the implications of equality of ratio of standard deviations. Firstly, we note that,

$$\frac{\sigma_{Y_1}}{\sigma_{X_1}} = \frac{\sigma_{Y_2}}{\sigma_{X_2}} \iff \frac{\sigma_{X_2}}{\sigma_{X_1}} = \frac{\sigma_{Y_2}}{\sigma_{Y_1}} \qquad (eqn\ ratios)$$

Secondly, using eqn ratios, we can argue as follows,

$$\frac{\sigma_{Y_1}}{\sigma_{X_1}} = \frac{\sigma_{Y_2}}{\sigma_{X_2}} \iff \left(\sigma_{Y_2} = \frac{\sigma_{X_2}}{\sigma_{X_1}}\,\sigma_{Y_1} \ \wedge\ \sigma_{X_2} = \frac{\sigma_{Y_2}}{\sigma_{Y_1}}\,\sigma_{X_1}\right) \iff \left(\sigma_{Y_2} = \frac{\sigma_{X_2}}{\sigma_{X_1}}\,\sigma_{Y_1} \ \wedge\ \sigma_{X_2} = \frac{\sigma_{X_2}}{\sigma_{X_1}}\,\sigma_{X_1}\right)$$

$$\implies \big(\exists d \in \mathbb{R} \cdot \sigma_{Y_2} = d.\sigma_{Y_1} \ \wedge\ \sigma_{X_2} = d.\sigma_{X_1}\big)$$

2) Using 4, the fact that $r(X_1,Y_1) = r(X_2,Y_2)$, the property just derived in part 1), with $d = \frac{\sigma_{X_2}}{\sigma_{X_1}}$ and rules of square roots, we can reason as follows,

$$r\big(X_2,(Y_2-X_2)\big) = \frac{\sigma_{Y_2}.r(X_2,Y_2) - \sigma_{X_2}.r(X_2,X_2)}{\sqrt{\sigma_{Y_2}^2 + \sigma_{X_2}^2 - 2.\sigma_{X_2}.\sigma_{Y_2}.r(X_2,Y_2)}} = \frac{\sigma_{Y_2}.r(X_1,Y_1) - \sigma_{X_2}.r(X_1,X_1)}{\sqrt{\sigma_{Y_2}^2 + \sigma_{X_2}^2 - 2.\sigma_{X_2}.\sigma_{Y_2}.r(X_1,Y_1)}}$$

$$= \frac{d.\sigma_{Y_1}.r(X_1,Y_1) - d.\sigma_{X_1}.r(X_1,X_1)}{\sqrt{d^2.\sigma_{Y_1}^2 + d^2.\sigma_{X_1}^2 - 2.d.\sigma_{X_1}.d.\sigma_{Y_1}.r(X_1,Y_1)}} = \frac{d.(\sigma_{Y_1}.r(X_1,Y_1) - \sigma_{X_1}.r(X_1,X_1))}{d.\sqrt{\sigma_{Y_1}^2 + \sigma_{X_1}^2 - 2.\sigma_{X_1}.\sigma_{Y_1}.r(X_1,Y_1)}}$$

$$= r\big(X_1,(Y_1-X_1)\big).$$

QED

**Proposition 8**

If $\Delta = Y - X$ and $p\Delta = X.\beta$, where $\beta \in \mathbb{R}$, then,

1) $r(p\Delta, \Delta) = sign(\beta).r(X, \Delta)$; and
2) $r(X + p\Delta, \mathrm{Y}) = sign(1 + \beta).r(X, \mathrm{Y})$.

**Proof**

Both results are easy consequences of proposition 5.

1) $r(p\Delta, \Delta) = r(X.\beta, \Delta) = sign(\beta).r(X, \Delta)$.

2) $r(X + p\Delta, \mathrm{Y}) = r((X + (X.\beta)), \mathrm{Y}) = r((X.(1 + \beta)), \mathrm{Y}) = sign(1 + \beta).r(X, \mathrm{Y}) = r(X, \mathrm{Y})$.

QED

## Main Findings

**Theorem 1:**

Since $X$ will be standardised, we can adapt the finding in proposition 4, to give us the key relationship we need,

$$r\big(X, (Y - X)\big) = \frac{\sigma_Y\,.r(X, Y) - \sigma_X}{\sqrt{\sigma_Y^2 + 1 - 2\,.\sigma_Y.r(X, Y)}} \qquad (\text{eqn } Imprint)$$

Note, this equation can be found in (Oldham, 1962), and also in (Tu *et al.*, 2005).

**Proof**

Immediate from proposition 4. QED

Theorem 1 shows clearly that $r(X, (Y - X))$ is fully defined by the correlations $r(X, Y)$ and $r(X, X)$, along with the variability of $Y$. The correlation of $X$ with itself, i.e. $r(X, X)$, is a prominent aspect of this equation, which drives its oddities. $r(X, X)$ reflects the coupling in the equation that arises because $X$ appears in both the terms being correlated in $r(X, (Y - X))$. $r(X, X)$ is of course a constant, i.e. 1 for any $X$, so in fact, $\sigma_Y$ and $r(X, Y)$, are the only variables; accordingly, their size determines the extent to which the imprint of $X$ in $Y - X$ drives $r(X, (Y - X))$.

This leads to the key observation that, as $\sigma_Y$ gets smaller, $r(X, (Y - X))$ tends towards $-r(X, X)$, which equals $-1$. In other words, as the variability of Y decreases, the imprint of $X$ becomes increasingly prominent. This is shown in the next theorem.

**Theorem 2**

$$r\big(X, (Y - X)\big) \longrightarrow -r(X, X) = -1, \qquad \text{as } \sigma_Y \longrightarrow 0$$

**Proof**

The right hand side of equation *Imprint*, has five constituent terms, two in the numerator and three in the denominator. Of these five, three are products with the standard deviation of $Y$, i.e. $\sigma_Y$. Assuming all else is constant, as $\sigma_Y$ reduces, the absolute value of each of these three terms reduces towards zero. The rate of reduction is different amongst the three, but they will all decrease. Accordingly, as $\sigma_Y$ decreases, $r\big(X, (Y - X)\big)$ becomes increasingly determined by the two terms not

involving $\sigma_Y$, and thus, it tends towards $-\frac{r(X,X)}{\sqrt{+1}} = -r(X,X) = -1$.
QED

## Equality of Residuals

An important finding of section 5 of the main text, is that the residuals resulting from regressing Y onto X are the same as regressing Y-X onto X. We show in this section, that this equality of residuals is necessarily the case.

We focus on the following two equations,

Eqn 1) $Y = \tilde{X}.\beta_1 + \varepsilon_1$

Eqn 2) $Y - X = \tilde{X}.\beta_2 + \varepsilon_2$

where $\tilde{X}$ is the $N \times 2$ matrix, with first column being $X$ and second being the $N \times 1$ vector of ones (which provides the intercept term); $\beta_1$ and $\beta_2$ are $2 \times 1$ vectors of parameters and $Y$, $X$, $\varepsilon_1$ and $\varepsilon_2$ are $N \times 1$ vectors. As in the rest of this document, $Y$ and $X$ are our (demeaned) initial and outcome variables, while $\varepsilon_1$ and $\varepsilon_2$ are our residual error terms.

**Proposition 9**

If we assume that $\beta_1$ and $\beta_2$ are fit with ordinary least squares, with $\varepsilon_1$ and $\varepsilon_2$ the associated residuals, then, $\varepsilon_1 = \varepsilon_2$.

**Proof**

Under ordinary least squares, the parameters are set as follows.

$$\beta_1 = (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T\,Y \quad \text{(Eqn 3)}$$

$$\beta_2 = (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T\,(Y - X) \quad \text{(Eqn 4)}$$

We start with the second of these, and using left distributivity of matrices, and then substituting Eqn 3, we obtain the following.

$$\beta_2 = (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T\,(Y - X) = (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T Y - (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T X = \beta_1 - (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T X$$

Using the fact that the variable $X$ is demeaned, we can now evaluate the main term here as follows,

$$\beta_2 = \beta_1 - (\tilde{X}^T\tilde{X})^{-1}\,\tilde{X}^T X = \beta_1 - \begin{pmatrix} X^2 & \Sigma X \\ \Sigma X & N \end{pmatrix}^{-1} \begin{pmatrix} X^2 \\ \Sigma X \end{pmatrix} = \beta_1 - \frac{1}{A}\begin{pmatrix} N & -\Sigma X \\ -\Sigma X & X^2 \end{pmatrix}\begin{pmatrix} X^2 \\ \Sigma X \end{pmatrix}$$

where $X^2$ is the dot product of $X$ with itself, $\Sigma X$ is the sum of the vector $X$, and $A = NX^2 - \Sigma X \Sigma X$ is the determinant of the matrix being inverted. From here we can derive the following,

$$\beta_2 = \beta_1 - \frac{1}{A}\begin{pmatrix} NX^2 - \Sigma X \Sigma X \\ -\Sigma X . X^2 + X^2 . \Sigma X \end{pmatrix} = \beta_1 - \frac{1}{A}\begin{pmatrix} A \\ 0 \end{pmatrix} = \beta_1 - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

We can then substitute this equality for $\beta_2$ in eqn 2 and re-arrange to obtain,

$$Y - X = \tilde{X}\beta_2 + \varepsilon_2 = \tilde{X}\left(\beta_1 - \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + \varepsilon_2 = \tilde{X}\beta_1 - X + \varepsilon_2$$

It follows straightforwardly from here that,

$$Y - \tilde{X}\beta_1 = \varepsilon_2$$

i.e. $\varepsilon_1 = \varepsilon_2$, as required. QED

Proposition 9 shows that the residuals resulting from fitting equations 1 and 2 will be the same. A consequence of this is that the error variability will be the same. As a result of this, the factor that determines whether more variance is explained when regressing $Y$ onto $X$ or when regressing $Y - X$ onto $X$, is the variance available to explain. That is, the relative variance of $Y$ and $Y - X$ drive the $R^2$ values of these two regressions. This then implicates the variance of $Y$ and $X$ and in fact their covariance (which impacts the variance of $Y - X$).

More precisely, we can state the following.

1) If $\sigma^2_{(Y-X)}$ is big relative to $\sigma^2_Y$, then regressing $Y - X$ onto $X$ will explain more variability than regressing $Y$ onto $X$.

2) If $\sigma^2_{(Y-X)}$ is small relative to $\sigma^2_Y$, then regressing $Y - X$ onto $X$ will explain less variability than regressing $Y$ onto $X$.

**Supplementary Appendix B: illustrating the relationship between the correlations**

```
% This function illustrates the relationship
function [r_XY,std_Y,r2,r3] = CheckEqn1()

noise = [0.01:0.01:1,2:100]; % controls r(X,Y)
scale = [0.01:0.01:1,2:100]; % controls sigma_Y/sigma_X
X = single(randn(1000,1));
for j=1:length(noise)
    Y = X + single(randn(1000,1).*noise(j)); %Y is X plus noise
    Y = zscore(Y); % then scale to X so the actual scaling is consistent
    for k=1:length(scale)
        Yl = Y.*scale(k); % rescale to control the variability ratio
        r_XY(j,k) = corr(X,Y); % calculate the correlation with outcomes

        r2(j,k) = corr(X,Yl-X); % calculate the correlation with change
        std_Y(j,k) = std(Yl)./std(X); % record the variability ratio
        r3(j,k) = eqn_r_X_XminusY(r_XY(j),std_Y(j,k)); % check Equation 1
    end
end

% display the resulting surface (Figure 1)
figure,surf(log(std_Y),r_XY,r3,'edgecolor','none')
lighting flat
l = light('Position',[50 100 100]);
l = light('Position',[50 100 -50]);
l = light('Position',[50 -100 -50]);
l = light('Position',[-50 -15 29]);
l = light('Position',[-50 -15 -29]);
l = light('Position',[-50 15 -29]);
l = light('Position',[50 15 -29]);
l = light('Position',[50 15 -50]);
shading interp
xlabel('log ( sigmaY / sigmaX )')
ylabel('r(X,Y)')
zlabel('r(X,Y-X)')

% confirm that equation 1 does actually match 'empirical' r(X,Y-X)
figure,scatter(r2(:),r3(:))
xlabel('Empirical coefficients')
ylabel('Derived coefficients')

end

% This function implements Equation 1
function res = eqn_r_X_XminusY(r_XY,std_Y)

res = (((r_XY.*std_Y) - 1) ./ sqrt(1 + (std_Y).^2 - (2*(r_XY.*std_Y))));

end
```