**Supplemental materials**

**1 Transformation of *change* equations (c.f. Introduction Section of Main Body)**

We show that all the *change* equations can be re-expressed as simple direct linear regression, what we call standard-form regression, bringing them all into the same model framework.

This mathematical formulation is inspired by the Fugl-Meyer (FM) scale for assessing upper limb deficits post-stroke (Prabhakaran et al., 2008). Specifically, denotes performance scores at time point 1 (i.e., *FM-initial* in the main body of the paper), denotes performance at time point 2 (i.e., *FM-end*), and is the top of the scale, which corresponds to “normal” performance on the scale ( in the FM scale).

We begin by defining the basic change equations, so-called because the dependent variable is the change from time-point 1 to time-point 2 (i.e., recovery). For simplicity of presentation, we do not include noise terms, but these could be added in the usual way.

**Definitions**

1) Classic *proportional to lost* *function* is defined as,

where .

2) *Proportional to spared* *function* is defined as,

where .

Please note that the *proportional to spared function* model is not the mathematical inverse of the *proportional to lost function* model. Indeed, models differ by the inclusion of the scale’s maximum *Max* and can thus not be transformed into one another.

3) *Constant* is defined as,

where .

These three change formulae generate different patterns in all but one case, which is when , the boundary condition in which and there is no recovery.

The standard-form linear regression equation, which directly relates and , rather than through change, is defined as,

where . We use lower case to distinguish coefficient and intercept when associated with standard-form regression and upper case when associated with the change formulations.

We now show that each change formulation can be rearranged into a standard-form regression.

**Proposition 1 (Proportional to Lost)**

Given that ,

**Proof**

Assume , then we can reason as follows,

but now, if we let and , we obtain,

which is as required.

Proposition 1 shows that once the *proportional to lost* equation is turned into the standard-form regression equation, as goes up (in the range zero to ), comes down, according to the formula . Thus, stronger proportional recovery patterns, i.e., with bigger ’s, lead to increased intercepts and reduced slopes in the standard-form regression.

The set of possible linear relationships that *proportional to lost* encompasses are shown in **Figure 2** (panel A change form and D standard-form) in the main body. This makes it clear that *proportional to lost* reflects a restricted range of the standard-form regression. Indeed, we can highlight two varieties of *proportional to lost*, both of which are sub-models of standard-form regression:

a) *Liberal proportional to lost*: standard-form regression in which and . This generalisation enables, for example, *constant* recovery to be mixed with *strong* *proportional to lost.*

b) *Strong proportional to lost*: the classic change formulation of *proportional to lost*, which (putting mathematical coupling aside) is equivalent to standard-form regression with and .

Both these forms require the slope to be less than or equal to the identity function (i.e., ). *Liberal* *proportional to lost*, though, does not constrain the relationship between slope and intercept, while *strong* *proportional to lost* does, making *strong* a sub-model of *liberal*.

Notably, the *liberal* and the *strong* *proportional to lost* formulations behave differently with regard to ceiling. That is, it is easy to show that, in the absence of noise, *strong proportional to lost* cannot generate a ceiling effect, i.e., values above . Specifically, *strong* *proportional to lost* can be expressed as,

where, for a given , the second summand on the right is constant and the first summand is maximal when . Thus, the maximum value that can be obtained is , which equals . In contrast, since and are independent in *liberal* *proportional to lost,* both can be set high in their given range, generating values above .

Now, we move to consider *proportional to spared* function.

**Proposition 2 (Proportional to Spared)**

Given that ,

**Proof**

Straightforward.

So, proposition 2 implies that *proportional to spared* can be expressed in standard-form regression, with an intercept of zero and a slope that is steeper than or equal to the identity, but less than 2. This range of possible linear relationships is depicted in **Figure 2** (panel B change formulation and E standard-form).

As was the case for *proportional to lost*, we can also identify a liberal form of *proportional to spared*.

a) *Liberal proportional to spared*: standard-form regression in which and.

b) *Strong proportional to spared*: the classic change formulation of *proportional to spared*, which (apart from mathematical coupling) is equivalent to standard-form regression with and .

Both these forms require the slope to be greater than or equal to the identity function, but less than 2. *Strong* *proportional to spared*, though, does not allow the intercept to increase above zero, while *liberal proportional to spared* does, making *strong* a sub-model of *liberal*.

As shown in **Figure 2** (panel E), *strong proportional to spared* generates ceiling effects, even without adding noise once (i.e., once ). These ceiling effects increase in severity as (or, indeed, ) gets bigger. Ceiling effects become even more severe if the intercept is allowed to be greater than zero, as supported by *liberal proportional to spared.*

Finally, we consider *constant* recovery (which is, of course, different to constant for standard-form regression, which would be ).

**Proposition 3 (Constant Recovery)**

Given that ,

**Proof**

Straightforward.

This reformulation makes clear that *constant* recovery generates the range of possible linear relationships shown in **Figure 2** (panels C and F). For bigger than zero, it also generates increasingly severe ceiling effects as increases, even without a noise term.

**2 Summary of Model Formulations (c.f. Introduction Section of Main Body)**

We employ several different model formulations in this paper. We summarise these formulations and where they are used here.

**Ground Truth Simulations (c.f. subsection “Synthetic data simulation experiments” in “Results” section)**

We generate fake, i.e., synthetic, data according to (constrained) standard-form regression models, of the form,

where the constraints are exactly as in main-body **Figure 2**, i.e., *proportional to lost* (panel D), *proportional to spared* (panel E), and *constant* recovery (panel F).

**Fitted to Ground Truth Simulation Data (c.f. subsection “Synthetic data simulation experiments” in “Results” section)**

In all cases, we fit the model

to the ground-truth simulation data; so we obtain estimates of slope, , and intercept, (along with their confidence intervals). Then, we assess the accuracy of fitting by comparing to the true slope and intercept; c.f. previous section. One reason for fitting standard-form, rather than change-form, models is that we wanted to fit to the model that generated the data. Generating data from change-form models does not enable a *Y* variable to be generated, hence the focus on standard-form.

**Fitted to Human Data, before taking subsets to reduce Ceiling (c.f. subsection “Bayesian posterior distributions for stroke recovery prediction” in “Results” section and Figure 4)**

This is the first fitting to human data. We fit two models:

1. (unconstrained) standard-form regression, i.e., ; c.f. main-body **Figure 4[A]**.
2. Typical change-form regression, i.e., ; c.f. main-body **Figure 4[B]**.

To fully illustrate how extreme the effect of *compression enhanced coupling* can be on the fitting of the change form model, we allowed corresponding flexibility to the change- and standard-form equations and incorporated an intercept () into the change-form model.

**Fitted to Human Data, after taking subsets to reduce Ceiling (c.f. subsection “Final model comparisons (on human data) in the subset of FM-initial 10-45” of “Results” section)**

We fitted the following models to the human data.

1. (unconstrained) standard-form regression, i.e., .
2. Change-form *proportional to lost*, i.e., .
3. Change-form *proportional to spared*, i.e., .
4. Change-form *constant* recovery, i.e., .

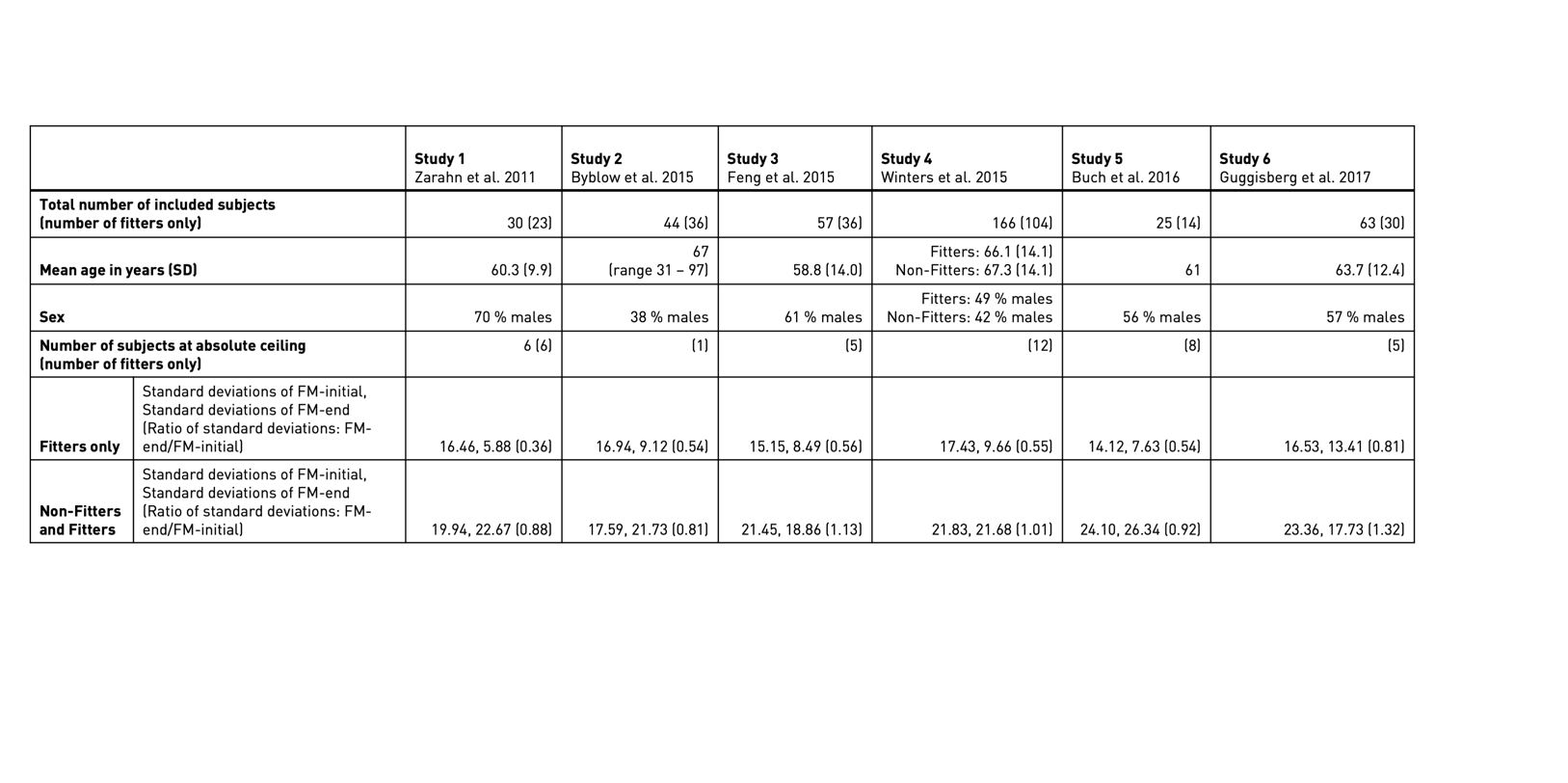
2, 3, and 4, here, are as per panels A, B, and C of main-body **Figure 2**.

We fitted the same models to the 0-45 subset of human data.

**3 Data acquisition and details on the Fugl-Meyer Scale (c.f. section Methods of Main Body)**

We contacted nine corresponding authors of eleven studies comprising longitudinal data on motor function measured on the Fugl-Meyer scale post-stroke (screening of titles & abstracts, keywords “poststroke”, “recovery”, “motor function”, “longitudinal”, “Fugl-Meyer” on PubMed as well as following referenced literature, criteria: > 20 stroke patients, initial and follow-up Fugl Meyer Score of the upper limb; initial: acute phase post-stroke, follow-up: 3-6 months post-stroke). The response rate was 56 %. Only two authors (Hawe, Scott, & Dukelow, 2019, Guggisberg, Nicolo, Cohen, Schnider, & Buch, 2017) were able to share their data (reasons for the decline: limited by ethics, in-house study). As individual-level data from (Zarahn et al., 2011) was openly available, we were, therefore, able to merge it with data from (Hawe et al., 2019) and (Guggisberg et al., 2017). Please see (Hawe et al., 2019) for a more detailed description of the data extraction. In brief, published figures of initial and end FM scores were digitized and positions of points extracted in Matlab, missing data was 17.7 %, primarily due to overlapping data points.

The FM assessment is a quantitative instrument to determine sensorimotor impairment post-stroke, divided into five domains (motor function, sensory function, balance, joint range of motion, and joint pain) (Fugl-Meyer *et al.*, 1975). Please note that we considered information on the motor performance of the affected upper limb only; a minimum of 0 implied no preserved and a maximum of 66 indicated full motor function of the upper limb (multiple items on movement, coordination, reflex action on the shoulder, elbow, forearm, wrist, hand, each evaluated on a 3-point ordinal scale (0 = no performance, 1 = partial performance, 2 = full performance).



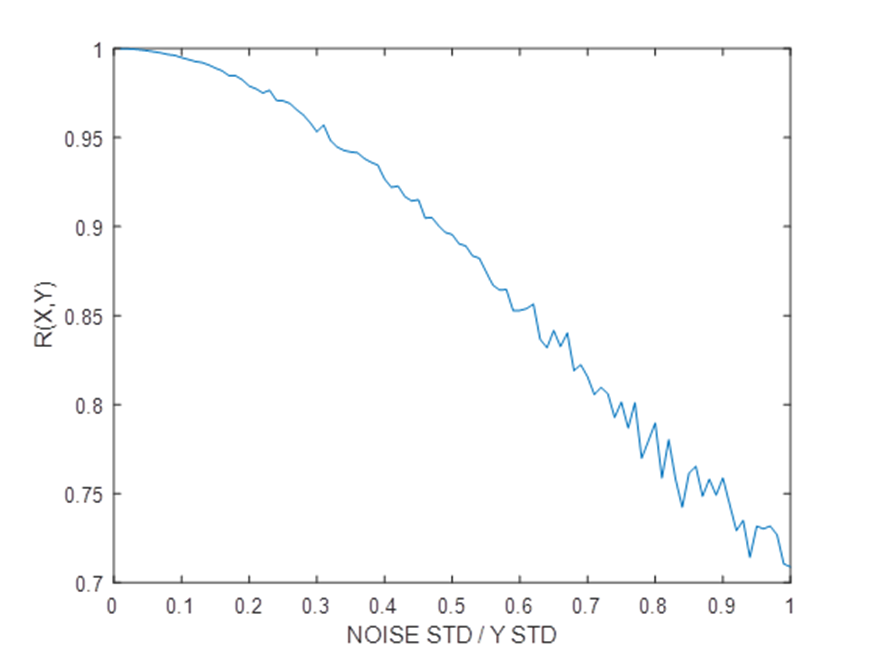
**Supplementary table 1. Descriptive statistics of each of the six studies included.**

**4 Defining simulated ceiling: Details on measurement noise (c.f. section Methods of Main Body)**

When simulating the outcome, *FM-end*, we first added “State noise” (Gaussian noise, mean = 0, Standard deviation (SD) = 5) to the output of the linear ground truth model. “State noise” can be understood as unexplained randomness associated with the internal state of the system, e.g., variation in the alertness of subjects. This form of noise reflects the actual capacity of the brain, ignoring the measurement method, and thus, naturally, is not bounded by the Fugl-Meyer scale*.*Next, we enforced ceiling by setting any value greater than 66 to 66, i.e., the maximum value of the Fugl-Meyer assessment. This imposition of ceiling reflects the application of the Fugl-Meyer scale and the parameters associated with it. This testing process will naturally involve a second source of random variability – so-called measurement noise. The latter reflects variability in the administration of the scale, which can be as subtle as the motivation a tester engenders in the patient and him- or herself. We justify our setting of measurement noise (Gaussian noise, mean = 0, SD = 2) in what follows. Since this further application of Gaussian noise will place some data points above ceiling, we enforced a final ceiling by once again setting any value greater than 66 to 66.

We relied on information about the test-retest reliability of the FM-scale, when estimating an appropriate level of measurement noise. Reported test-retest and inter-rater reliabilities for the Fugl-Meyer are *r>=0.96* over repeated tests (Gladstone, Danells, & Black, 2002). Accordingly, we assessed the ratio of the data SD that the noise SD has to be to obtain such an *r*-value. This correlation is between a first and second test, where the only difference between the two is the addition of the noise.

The results of this simulation are shown in **Supplementary Figure 1,** which demonstrates that noise with 10-20% (or even up to 30%) of the standard deviation of the original scores, gives strong enough correlations to match empirically reported *r-*values. Given our “true” models, standard deviations of the outcome had a range of 5.3 – 33.7. For reasons of simplicity, we calculated a probable level of noise based on the smallest occurring standard deviation, i.e., SD = 5.3, and an assumption of 30 % noise for all models; this gives us a rounded measurement noise standard deviation value of 2.



**Supplementary Figure 1.** Correlation of variables X and Y, representing two measurements, where X is Y plus zero mean Gaussian noise. The standard deviation of the Gaussian noise is varied relative to a fixed standard deviation for X.

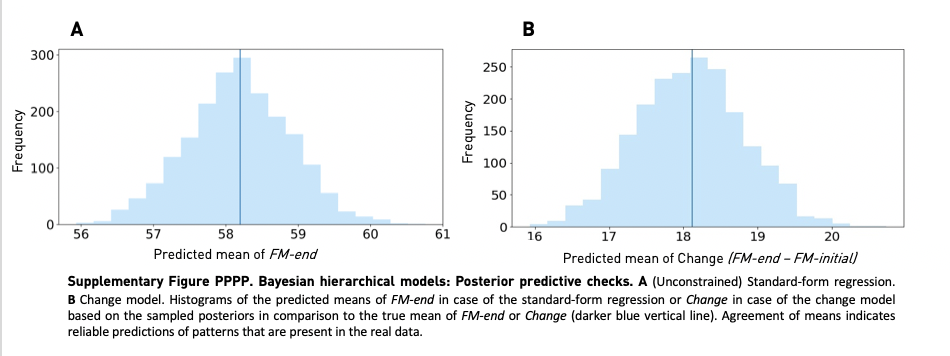
**5 Additional comments on simulations (c.f. section Methods of Main Body)**

Importantly, the model comparison problem we are undertaking is simplified because the linear fits provided by each of the three recovery theories – *proportional to lost, to spared,* and *constant* – are, in all but one case, non-overlapping. This overlap can be seen in main-body panels [A. B & C], as well as [D, E & F] of **Figure 2**. Specifically, whether in change or standard-form, apart from the identity mapping, i.e., Y = X (which all the models allow, but will not arise in our fitting), the model spaces are disjoint. This effectively ensures *model* retrievability, i.e., if we can show that each model can retrieve the parameters of synthetic data it generated (the issue considered in section 7 of the supplementary material), no non-generating model will fit better than a generating model.

Of course, this only applies to the three specific recovery models – *proportional to lost, proportional to spared,* and *constant* recovery. We will also fit (unconstrained) standard-form regression, which can generate all the linear fits of the three specific models, as well as fits that they cannot generate.

To increase the accessibility of our simulations, we illustrate a typical simulation process in **Figure 1[A]** of the main manuscript**:**The black cross marks data generated from a ground truth model with known properties and initially without ceiling enforced. The position the cross marks has *r(X,Y-X)* and *r(X,Y)* as strong positive correlations and similar variabilities for *X* and *Y* (i.e., *variability ratio* around one, its log around zero). This, then, is synthetic data consistent with a *proportional to spared pattern* of recovery (**Figure 2[B,E])**. We then simulate the effect of ceiling, thereby, reducing the *variability ratio*, which, in turn, causes the data to exhibit a *proportional to lost pattern*, i.e., *r(X,Y-X)* becomes negative, and it moves in the surface plot (c.f., black dashed arrow) towards the flat region (c.f., black star). In particular, we will be able to show that by employing a subset to remove ceiling (**Figure 1[C,D]**), we can return the synthetic data, via dotted black arrow, to ground truth position, i.e., black cross.

**6 Bayesian hierarchical models for the standard-form regression and change model of *fitters* (*FM-initial*: > 10): Posterior predictive checks (c.f. Results section of Main Body)**

****

**Supplementary Figure 2. Bayesian hierarchical models: Posterior predictive checks. A** (Unconstrained) Standard-form regression. **B** Change model. Histograms of the predicted means of *FM-end* in case of the standard-form regression or *Change* in case of the change model, based on the sampled posteriors in comparison to the true mean of *FM-end* or *Change* (darker blue vertical line). Agreement of means indicates reliable predictions of patterns that are present in the real data.

**7 Detailed description and evaluation of simulations (c.f. Methods and Results section of Main Body)**

We report here the results of our ground truth simulations, which are performed for the three key models: *proportional to lost function* recovery, *proportional to spared function* recovery, and *constant* recovery. In each of these cases, data is generated according to the standard-form regression version of each model, i.e., the equations presented in panels [D, E, and F] of **Figure 2**. These re-expressions are justified in **supplementary materials section 1**. Thus, the generating models can be expressed as *Y=b.X+a*, with *b* the true slope and *a* the true intercept in the tables. Notably, the resulting variances explained would not typically be the same, since only the change formulation exhibits mathematical coupling; for a relevant discussion, c.f., section Equality of Variances in Supplementary Appendix A of (Hope *et al.*, 2019). Also, for *proportional to lost* function recovery *b=1-B* and *a=B.Max*, with the “amount of recovery” (1st column of tables) equal to *B.100*; for *proportional to spared function* recovery *b=(1+B)* and *a=0*, with “amount of recovery” equal to *B.100*; and for *constant* recovery *b=1* and *a=A*, with “amount of recovery” equal to *A*.

Each row in each table below corresponds to simulations in which we generate 1000 data sets according to given ground truth (i.e., setting of *b* and *a*) and then fit the same model (*Y=b.X+a*) to each data set. The critical criterion for judging the effectiveness of model retrieval is the number of times the “correct” (ground truth) parameter settings (i.e., for *a* and *b*) were in the confidence intervals resulting from model fitting; we report the proportion that met the 68% and 95% Confidence Intervals (CIs) (Gelman & Hill, 2006).

Model retrieval is maximally accurate if 68% (respectively 95%) of fits are in the 68% (respectively 95%) CIs. We also report the variability ratio, i.e., the ratio of standard deviations of *FM-end (Y)* and *FM-initial* (*X)* (log of this number is shown in main-body **Figure 1[A]**) and then the two relevant correlations: *r(X,Y)* and *r(X,Y-X*), i.e., *FM-initial – FM-end* and *FM-initial – Change*. Finally, we report the average number of subjects at or above absolute ceiling, i.e., the . We are referring to “absolute” ceiling as further subjects might present with confounded scores that are compressed toward ceiling, yet not at *FM-end*=66.

We first state a summary of simulation results and then present the corresponding tables:

**Proportional to lost function recovery**

**Baseline models (state noise only):** *Proportional to lost function* recovery in the range of 10 % to 90 % was characterized by positive intercepts and slopes in-between values of 0 and 1. While intercepts increased with the amount of *proportional to lost function* recovery (range: 6.6 – 59.4), the opposite behavior was found for slopes (range: 0.9 – 0.2), **Supplementary Table 1a** illustrates these trajectories. This is consistent with *proportional to lost function* recovery when formulated as a standard-form regression; see **proposition 1** in supplementary materials (part 1) and main-body **Figure 2** panel [D]. Additionally, as expected and highlighted in (Hope *et al.*, 2018) and (Hawe *et al.*, 2019*a*), Pearson correlation values steadily decreased for *FM-initial – Change* until an almost perfect anti-correlation of -1, which was contrasted with decreasing positive correlation values for *FM-end – FM-initial*, when increasing absolute percentages of *proportional to lost function* recovery. Further, higher percentages of *proportional to lost function* were associated with a decrease in *variability ratio*, reaching a minimum value of 0.3 for 90% recovery (c.f. main-body **Figure 6**, A and B, left plots). In case of 70 % recovery, the absolute value of the correlation of *FM-initial – Change* surpassed *FM-end – FM-initial* by 0.24, therefore, perfectly demonstrating the inflation of correlation due to mathematical coupling and its enhancement as the *variability ratio* reduces.

**Ceiling enforced:** Enforcing ceiling aggravated the dissociation of correlations of *FM-initial – Change* and *FM-end – FM-initial*, in line with predictions by (Hope *et al.*, 2018). The percentage of subjects at absolute ceiling (i.e. FM-end = 66) reached a maximum of 21.4 % for 90 % recovery (range: 4.1 % - 21.4 %, 70 % proportional recovery: 14.8 %).

**Subset approaches FM-initial 10 to 60, 50 and 40:** Reducing the number of subjects according to their *FM-initial* scores re-established accurate inferences of intercepts and slopes, as was indicated by met confidence intervals. The most substantial subset *FM-initial* 10 to 60 enabled veridical model estimation up to 30 % of *proportional to lost function* recovery, *FM-initial* 10 to 50 up to 70 % and *FM-initial* 10 to 40 for the entirety of models. This increase in accurate model estimation was paralleled by a decrease in subjects at ceiling.

**Supplementary table 1a: Proportional to lost function recovery: No ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 6.61 | 0.69 | 0.95 | 0.90 | 0.90 | 0.70 | 0.94 |
| 20% | 13.20 | 13.21 | 0.68 | 0.95 | 0.80 | 0.80 | 0.71 | 0.96 |
| 30% | 19.80 | 19.81 | 0.69 | 0.95 | 0.70 | 0.70 | 0.69 | 0.96 |
| 40% | 26.40 | 26.38 | 0.68 | 0.96 | 0.60 | 0.60 | 0.68 | 0.94 |
| 50% | 33.00 | 32.94 | 0.70 | 0.96 | 0.50 | 0.50 | 0.68 | 0.96 |
| 60% | 39.60 | 39.57 | 0.68 | 0.95 | 0.40 | 0.40 | 0.70 | 0.94 |
| 70% | 46.20 | 46.20 | 0.67 | 0.94 | 0.30 | 0.30 | 0.69 | 0.92 |
| 80% | 52.80 | 52.83 | 0.69 | 0.95 | 0.20 | 0.20 | 0.69 | 0.95 |
| 90% | 59.40 | 59.37 | 0.71 | 0.94 | 0.10 | 0.10 | 0.68 | 0.95 |

|  |  |  |
| --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** |
| 0.94 | 0.95 | -0.33 |
| 0.85 | 0.94 | -0.57 |
| 0.76 | 0.93 | -0.72 |
| 0.66 | 0.90 | -0.81 |
| 0.58 | 0.87 | -0.87 |
| 0.49 | 0.82 | -0.90 |
| 0.41 | 0.72 | -0.93 |
| 0.35 | 0.57 | -0.94 |
| 0.30 | 0.33 | -0.95 |

**Supplementary table 1b: Proportional to lost function recovery: Ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 7.09 | 0.61 | 0.90 | 0.90 | 0.88 | 0.49 | 0.85 |
| 20% | 13.20 | 13.72 | 0.59 | 0.90 | 0.80 | 0.78 | 0.46 | 0.83 |
| 30% | 19.80 | 20.38 | 0.56 | 0.87 | 0.70 | 0.68 | 0.43 | 0.80 |
| 40% | 26.40 | 27.09 | 0.51 | 0.83 | 0.60 | 0.57 | 0.35 | 0.70 |
| 50% | 33.00 | 33.71 | 0.51 | 0.85 | 0.50 | 0.47 | 0.27 | 0.68 |
| 60% | 39.60 | 40.37 | 0.47 | 0.82 | 0.40 | 0.37 | 0.22 | 0.55 |
| 70% | 46.20 | 47.01 | 0.42 | 0.79 | 0.30 | 0.26 | 0.14 | 0.44 |
| 80% | 52.80 | 53.56 | 0.42 | 0.81 | 0.20 | 0.16 | 0.06 | 0.28 |
| 90% | 59.40 | 59.39 | 0.61 | 0.92 | 0.10 | 0.07 | 0.08 | 0.33 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 0.93 | 0.95 | -0.37 | 10 |
| 0.83 | 0.94 | -0.60 | 11 |
| 0.74 | 0.92 | -0.74 | 13 |
| 0.64 | 0.90 | -0.83 | 14 |
| 0.55 | 0.86 | -0.88 | 17 |
| 0.46 | 0.80 | -0.92 | 21 |
| 0.38 | 0.70 | -0.94 | 26 |
| 0.30 | 0.53 | -0.96 | 35 |
| 0.23 | 0.29 | -0.97 | 52 |

**Supplementary table 1c: Proportional to lost function recovery: Ceiling enforced, Subset 10 - 60.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 6.68 | 0.67 | 0.94 | 0.90 | 0.90 | 0.69 | 0.95 |
| 20% | 13.20 | 13.29 | 0.69 | 0.94 | 0.80 | 0.80 | 0.68 | 0.94 |
| 30% | 19.80 | 19.95 | 0.66 | 0.94 | 0.70 | 0.69 | 0.67 | 0.94 |
| 40% | 26.40 | 26.62 | 0.67 | 0.95 | 0.60 | 0.59 | 0.65 | 0.93 |
| 50% | 33.00 | 33.32 | 0.64 | 0.92 | 0.50 | 0.49 | 0.60 | 0.91 |
| 60% | 39.60 | 40.07 | 0.61 | 0.91 | 0.40 | 0.38 | 0.54 | 0.86 |
| 70% | 46.20 | 46.76 | 0.57 | 0.89 | 0.30 | 0.27 | 0.40 | 0.77 |
| 80% | 52.80 | 53.38 | 0.52 | 0.86 | 0.20 | 0.17 | 0.25 | 0.61 |
| 90% | 59.40 | 59.33 | 0.65 | 0.93 | 0.10 | 0.07 | 0.20 | 0.55 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 0.96 | 0.94 | -0.29 | 2 |
| 0.86 | 0.92 | -0.52 | 3 |
| 0.77 | 0.90 | -0.68 | 4 |
| 0.68 | 0.87 | -0.78 | 5 |
| 0.58 | 0.83 | -0.85 | 8 |
| 0.50 | 0.77 | -0.89 | 11 |
| 0.41 | 0.66 | -0.92 | 16 |
| 0.33 | 0.50 | -0.94 | 24 |
| 0.26 | 0.26 | -0.97 | 41 |

**Supplementary table 1d: Proportional to lost function recovery: Ceiling enforced, Subset 10 - 50.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 6.64 | 0.66 | 0.95 | 0.90 | 0.90 | 0.67 | 0.95 |
| 20% | 13.20 | 13.14 | 0.69 | 0.95 | 0.80 | 0.80 | 0.69 | 0.94 |
| 30% | 19.80 | 19.86 | 0.69 | 0.96 | 0.70 | 0.70 | 0.70 | 0.96 |
| 40% | 26.40 | 26.42 | 0.65 | 0.94 | 0.60 | 0.60 | 0.67 | 0.95 |
| 50% | 33.00 | 33.02 | 0.67 | 0.94 | 0.50 | 0.50 | 0.67 | 0.95 |
| 60% | 39.60 | 39.78 | 0.69 | 0.95 | 0.40 | 0.39 | 0.68 | 0.95 |
| 70% | 46.20 | 46.43 | 0.65 | 0.95 | 0.30 | 0.29 | 0.69 | 0.95 |
| 80% | 52.80 | 53.13 | 0.62 | 0.92 | 0.20 | 0.18 | 0.55 | 0.90 |
| 90% | 59.40 | 59.27 | 0.67 | 0.94 | 0.10 | 0.07 | 0.46 | 0.81 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 0.99 | 0.90 | -0.23 | 0 |
| 0.91 | 0.88 | -0.42 | 0 |
| 0.82 | 0.85 | -0.58 | 0 |
| 0.73 | 0.82 | -0.69 | 1 |
| 0.65 | 0.76 | -0.76 | 1 |
| 0.57 | 0.68 | -0.82 | 3 |
| 0.50 | 0.58 | -0.87 | 6 |
| 0.42 | 0.42 | -0.91 | 12 |
| 0.34 | 0.21 | -0.94 | 27 |

**Supplementary table 1e: Proportional to lost function recovery: Ceiling enforced, Subset 10 - 40.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 6.58 | 0.68 | 0.95 | 0.90 | 0.90 | 0.68 | 0.95 |
| 20% | 13.20 | 13.13 | 0.69 | 0.95 | 0.80 | 0.80 | 0.67 | 0.94 |
| 30% | 19.80 | 19.78 | 0.68 | 0.95 | 0.70 | 0.70 | 0.68 | 0.95 |
| 40% | 26.40 | 26.37 | 0.68 | 0.95 | 0.60 | 0.60 | 0.68 | 0.95 |
| 50% | 33.00 | 33.06 | 0.66 | 0.95 | 0.50 | 0.50 | 0.67 | 0.94 |
| 60% | 39.60 | 39.68 | 0.69 | 0.94 | 0.40 | 0.40 | 0.69 | 0.94 |
| 70% | 46.20 | 46.28 | 0.68 | 0.96 | 0.30 | 0.29 | 0.71 | 0.96 |
| 80% | 52.80 | 52.98 | 0.68 | 0.95 | 0.20 | 0.18 | 0.69 | 0.95 |
| 90% | 59.40 | 59.17 | 0.64 | 0.94 | 0.10 | 0.08 | 0.64 | 0.91 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.08 | 0.83 | -0.17 | 0.00 |
| 1.00 | 0.80 | -0.31 | 0.00 |
| 0.92 | 0.76 | -0.45 | 0.00 |
| 0.85 | 0.71 | -0.55 | 0.00 |
| 0.78 | 0.64 | -0.64 | 0.00 |
| 0.71 | 0.55 | -0.71 | 0.00 |
| 0.66 | 0.45 | -0.77 | 1.00 |
| 0.59 | 0.31 | -0.82 | 5.00 |
| 0.49 | 0.15 | -0.89 | 14.00 |

**Supplementary table 1f: Proportional to lost function recovery: Ceiling enforced, Subset 10 - 45.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 6.60 | 6.54 | 0.68 | 0.94 | 0.90 | 0.90 | 0.69 | 0.95 |
| 20% | 13.20 | 13.30 | 0.68 | 0.95 | 0.80 | 0.80 | 0.69 | 0.94 |
| 30% | 19.80 | 19.89 | 0.68 | 0.96 | 0.70 | 0.70 | 0.69 | 0.95 |
| 40% | 26.40 | 26.36 | 0.68 | 0.94 | 0.60 | 0.60 | 0.67 | 0.94 |
| 50% | 33.00 | 33.01 | 0.68 | 0.95 | 0.50 | 0.50 | 0.67 | 0.95 |
| 60% | 39.60 | 39.63 | 0.66 | 0.96 | 0.40 | 0.40 | 0.67 | 0.95 |
| 70% | 46.20 | 46.33 | 0.67 | 0.94 | 0.30 | 0.29 | 0.66 | 0.94 |
| 80% | 52.80 | 53.06 | 0.67 | 0.94 | 0.20 | 0.18 | 0.65 | 0.93 |
| 90% | 59.40 | 59.24 | 0.67 | 0.93 | 0.10 | 0.07 | 0.55 | 0.89 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.03 | 0.87 | -0.19 | 0.00 |
| 0.94 | 0.85 | -0.37 | 0.00 |
| 0.86 | 0.81 | -0.52 | 0.00 |
| 0.78 | 0.77 | -0.62 | 0.00 |
| 0.71 | 0.70 | -0.70 | 0.00 |
| 0.64 | 0.62 | -0.77 | 1.00 |
| 0.57 | 0.51 | -0.82 | 3.00 |
| 0.50 | 0.36 | -0.87 | 7.00 |
| 0.40 | 0.18 | -0.92 | 19.00 |

**Proportional to spared function recovery**

**Baseline models (state noise only):** It is important to note that – in contrast *to proportional to lost function* – when formulated in standard-form, *proportional to spared function* is characterized by intercepts of zero and slopes greater than one (and less than 2); see main-body **Figure 2, panel [E].** Pearson correlations of *FM-initial – Change* as well as *FM-end – FM-initial* reached maximal values of almost 1 for increasing degrees of *proportional to spared function* recovery. These were accompanied by variability ratios (of standard deviations) greater than one (90 %: 1.9) (c.f. main-body **Figure 6, [A&B],** center plots). This is the regime “on the hill” in our surface plot, i.e., log variability ratio substantially greater than zero, main-body **Figure 1**.

**Ceiling enforced:** Setting any *FM-end* values greater than it to 66 had a detrimental effect on the accurate estimation of parameters for all *proportional to spared* instantiations, confidence intervals were not met in any of the cases. Crucially, estimated intercept and slope combinations resembled those of the true *proportional to lost* function and *constant* recovery models: Intercepts increased in parallel to increasing amounts of proportional to spared function, while slopes started at values close to one and successively decreased to a minimum of 0.76. Even in the case of only 10 % *proportional to spared function*, there were 9.1 % of subjects at absolute ceiling (maximum: 90 % *proportional to spared function* recovery with 30.9 % subjects at ceiling).

**Subset approaches FM-initial 10 to 60, 50, 40:** Once again, accuracy of estimation performance increased with decreasing upper bounds for the FM-initial scores, yet not as completely as for *proportional to lost function*: *FM-initial* 10 – 60: up to 10 %, *FM-initial* 10 – 50: up to 30 % and *FM-initial* 10 – 40: up to 70 % of *proportional to spared function* recovery.

**Supplementary table 2a: Proportional to spared function recovery: No ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 0.00 | 0.67 | 0.94 | 1.10 | 1.10 | 0.71 | 0.94 |
| 20% | 0.00 | 0.01 | 0.70 | 0.95 | 1.20 | 1.20 | 0.68 | 0.95 |
| 30% | 0.00 | 0.01 | 0.71 | 0.95 | 1.30 | 1.30 | 0.71 | 0.95 |
| 40% | 0.00 | -0.01 | 0.67 | 0.95 | 1.40 | 1.40 | 0.68 | 0.96 |
| 50% | 0.00 | 0.00 | 0.67 | 0.95 | 1.50 | 1.50 | 0.67 | 0.94 |
| 60% | 0.00 | 0.00 | 0.69 | 0.95 | 1.60 | 1.60 | 0.69 | 0.95 |
| 70% | 0.00 | -0.02 | 0.71 | 0.96 | 1.70 | 1.70 | 0.68 | 0.96 |
| 80% | 0.00 | -0.02 | 0.67 | 0.94 | 1.80 | 1.80 | 0.68 | 0.94 |
| 90% | 0.00 | 0.00 | 0.70 | 0.96 | 1.90 | 1.90 | 0.68 | 0.96 |

|  |  |  |
| --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** |
| 1.11 | 0.99 | 0.66 |
| 1.21 | 1.00 | 0.87 |
| 1.30 | 1.00 | 0.93 |
| 1.40 | 1.00 | 0.96 |
| 1.50 | 1.00 | 0.98 |
| 1.60 | 1.00 | 0.98 |
| 1.70 | 1.00 | 0.99 |
| 1.80 | 1.00 | 0.99 |
| 1.90 | 1.00 | 0.99 |

**Supplementary table 2b: Proportional to spared function recovery: Ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 1.80 | 0.12 | 0.39 | 1.10 | 1.03 | 0.00 | 0.03 |
| 20% | 0.00 | 4.17 | 0.00 | 0.00 | 1.20 | 1.04 | 0.00 | 0.00 |
| 30% | 0.00 | 7.02 | 0.00 | 0.00 | 1.30 | 1.02 | 0.00 | 0.00 |
| 40% | 0.00 | 10.06 | 0.00 | 0.00 | 1.40 | 0.99 | 0.00 | 0.00 |
| 50% | 0.00 | 13.21 | 0.00 | 0.00 | 1.50 | 0.94 | 0.00 | 0.00 |
| 60% | 0.00 | 16.26 | 0.00 | 0.00 | 1.60 | 0.90 | 0.00 | 0.00 |
| 70% | 0.00 | 19.20 | 0.00 | 0.00 | 1.70 | 0.85 | 0.00 | 0.00 |
| 80% | 0.00 | 22.05 | 0.00 | 0.00 | 1.80 | 0.80 | 0.00 | 0.00 |
| 90% | 0.00 | 24.68 | 0.00 | 0.00 | 1.90 | 0.76 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.07 | 0.96 | 0.11 | 22 |
| 1.08 | 0.96 | 0.13 | 36 |
| 1.07 | 0.95 | 0.06 | 46 |
| 1.05 | 0.94 | -0.04 | 54 |
| 1.02 | 0.92 | -0.15 | 61 |
| 0.99 | 0.91 | -0.25 | 65 |
| 0.95 | 0.89 | -0.33 | 68 |
| 0.92 | 0.87 | -0.41 | 72 |
| 0.88 | 0.86 | -0.47 | 75 |

**Supplementary table 2c: Proportional to spared function recovery: Ceiling enforced, Subset 10 - 60.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 0.46 | 0.61 | 0.92 | 1.10 | 1.08 | 0.54 | 0.87 |
| 20% | 0.00 | 1.73 | 0.17 | 0.50 | 1.20 | 1.13 | 0.00 | 0.08 |
| 30% | 0.00 | 3.75 | 0.00 | 0.01 | 1.30 | 1.14 | 0.00 | 0.00 |
| 40% | 0.00 | 6.42 | 0.00 | 0.00 | 1.40 | 1.12 | 0.00 | 0.00 |
| 50% | 0.00 | 9.30 | 0.00 | 0.00 | 1.50 | 1.09 | 0.00 | 0.00 |
| 60% | 0.00 | 12.26 | 0.00 | 0.00 | 1.60 | 1.04 | 0.00 | 0.00 |
| 70% | 0.00 | 15.24 | 0.00 | 0.00 | 1.70 | 1.00 | 0.00 | 0.00 |
| 80% | 0.00 | 18.09 | 0.00 | 0.00 | 1.80 | 0.95 | 0.00 | 0.00 |
| 90% | 0.00 | 20.82 | 0.00 | 0.00 | 1.90 | 0.90 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.13 | 0.96 | 0.24 | 8 |
| 1.17 | 0.96 | 0.38 | 18 |
| 1.19 | 0.96 | 0.39 | 28 |
| 1.18 | 0.95 | 0.32 | 36 |
| 1.15 | 0.94 | 0.22 | 42 |
| 1.12 | 0.93 | 0.11 | 46 |
| 1.09 | 0.91 | -0.01 | 49 |
| 1.06 | 0.90 | -0.11 | 53 |
| 1.02 | 0.88 | -0.20 | 56 |

**Supplementary table 2d: Proportional to spared function recovery: Ceiling enforced, Subset 10 - 50.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 0.04 | 0.69 | 0.95 | 1.10 | 1.10 | 0.66 | 0.95 |
| 20% | 0.00 | 0.14 | 0.67 | 0.95 | 1.20 | 1.19 | 0.68 | 0.94 |
| 30% | 0.00 | 0.38 | 0.69 | 0.93 | 1.30 | 1.28 | 0.64 | 0.92 |
| 40% | 0.00 | 1.31 | 0.37 | 0.75 | 1.40 | 1.34 | 0.15 | 0.51 |
| 50% | 0.00 | 3.08 | 0.03 | 0.18 | 1.50 | 1.35 | 0.00 | 0.00 |
| 60% | 0.00 | 5.24 | 0.00 | 0.00 | 1.60 | 1.34 | 0.00 | 0.00 |
| 70% | 0.00 | 7.72 | 0.00 | 0.00 | 1.70 | 1.31 | 0.00 | 0.00 |
| 80% | 0.00 | 10.40 | 0.00 | 0.00 | 1.80 | 1.28 | 0.00 | 0.00 |
| 90% | 0.00 | 13.02 | 0.00 | 0.00 | 1.90 | 1.23 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.18 | 0.93 | 0.22 | 0 |
| 1.27 | 0.94 | 0.42 | 1 |
| 1.35 | 0.95 | 0.56 | 5 |
| 1.39 | 0.96 | 0.64 | 10 |
| 1.41 | 0.96 | 0.66 | 16 |
| 1.41 | 0.95 | 0.62 | 20 |
| 1.39 | 0.94 | 0.57 | 23 |
| 1.37 | 0.93 | 0.49 | 27 |
| 1.34 | 0.92 | 0.41 | 30 |

**Supplementary table 2e: Proportional to spared function recovery: Ceiling enforced, Subset 10 - 40.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 0.02 | 0.71 | 0.95 | 1.10 | 1.10 | 0.68 | 0.96 |
| 20% | 0.00 | 0.01 | 0.69 | 0.96 | 1.20 | 1.20 | 0.69 | 0.95 |
| 30% | 0.00 | 0.00 | 0.67 | 0.95 | 1.30 | 1.30 | 0.68 | 0.95 |
| 40% | 0.00 | -0.04 | 0.70 | 0.95 | 1.40 | 1.40 | 0.70 | 0.95 |
| 50% | 0.00 | 0.02 | 0.68 | 0.95 | 1.50 | 1.50 | 0.67 | 0.95 |
| 60% | 0.00 | 0.22 | 0.66 | 0.95 | 1.60 | 1.59 | 0.68 | 0.95 |
| 70% | 0.00 | 0.61 | 0.65 | 0.93 | 1.70 | 1.67 | 0.63 | 0.93 |
| 80% | 0.00 | 1.36 | 0.49 | 0.86 | 1.80 | 1.72 | 0.34 | 0.76 |
| 90% | 0.00 | 2.71 | 0.19 | 0.55 | 1.90 | 1.74 | 0.03 | 0.18 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.25 | 0.88 | 0.16 | 0.00 |
| 1.34 | 0.90 | 0.32 | 0.00 |
| 1.43 | 0.91 | 0.45 | 0.00 |
| 1.52 | 0.92 | 0.56 | 0.00 |
| 1.61 | 0.93 | 0.64 | 0.00 |
| 1.69 | 0.94 | 0.71 | 1.00 |
| 1.76 | 0.94 | 0.75 | 3.00 |
| 1.81 | 0.95 | 0.79 | 6.00 |
| 1.83 | 0.95 | 0.80 | 8.00 |

**Supplementary table 2f: Proportional to spared function recovery: Ceiling enforced, Subset 10 - 45.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 10% | 0.00 | 0.06 | 0.69 | 0.95 | 1.10 | 1.10 | 0.67 | 0.94 |
| 20% | 0.00 | -0.01 | 0.67 | 0.94 | 1.20 | 1.20 | 0.67 | 0.94 |
| 30% | 0.00 | 0.07 | 0.71 | 0.95 | 1.30 | 1.30 | 0.69 | 0.94 |
| 40% | 0.00 | 0.24 | 0.71 | 0.95 | 1.40 | 1.39 | 0.70 | 0.96 |
| 50% | 0.00 | 0.78 | 0.61 | 0.89 | 1.50 | 1.46 | 0.54 | 0.88 |
| 60% | 0.00 | 1.94 | 0.28 | 0.65 | 1.60 | 1.50 | 0.08 | 0.38 |
| 70% | 0.00 | 3.36 | 0.04 | 0.24 | 1.70 | 1.52 | 0.00 | 0.01 |
| 80% | 0.00 | 5.33 | 0.00 | 0.01 | 1.80 | 1.51 | 0.00 | 0.00 |
| 90% | 0.00 | 7.56 | 0.00 | 0.00 | 1.90 | 1.49 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.21 | 0.91 | 0.19 | 0.00 |
| 1.30 | 0.92 | 0.37 | 0.00 |
| 1.39 | 0.93 | 0.51 | 0.00 |
| 1.47 | 0.94 | 0.62 | 2.00 |
| 1.54 | 0.95 | 0.69 | 4.00 |
| 1.58 | 0.95 | 0.72 | 7.00 |
| 1.60 | 0.95 | 0.72 | 10.00 |
| 1.60 | 0.95 | 0.70 | 13.00 |
| 1.59 | 0.94 | 0.67 | 16.00 |

**Constant recovery**

**Baseline models (state noise only):** When formulated in standard-form regression, the natural behavior of *constant* recovery models is expressed by increasing intercepts, with unvarying slopes at one; see **Figure 2[F].** Pearson correlations for*FM-end – FM-initial,* as well as the ratio of standard deviations, occupy values close to 1, *FM-initial – Change,* in contrast, are not correlated; these phenomena are independent of the amount of *constant* recovery (c.f. **Figure 6, [A&B]**, right plots).

**Ceiling enforced:** The ceiling effect rendered it impossible to veridically estimate the true model parameters. Confidence intervals were not met in any simulation run for any *constant* model above 5 points recovery. Of further note, the variability ratio of standard deviations decreased from 0.99 to a minimum of 0.13; see **Figure 6 [B],** rightmost panel, thus becoming even more extreme than the ratio for *proportional to lost* function. A dissociation of correlations also reflected this: the correlation of*FM-initial – Change* dropped from zero to an almost perfect anticorrelation of -1, while the correlation of *FM-end – FM-initial* showed (in the positive direction) the opposite behavior, starting at 0.96, yet dropping to 0.36 for a maximum of 50 points *constant* recovery; see **Figure 6 [A]** rightmost panel. The percentages of subjects at ceiling increased from 8.2 % to 45.3 %. The numbers of subjects at ceiling for a *constant* recovery of 50 points (n=110, out of 243, 45%) was also the highest observed across all recovery models.

**Subset approaches FM-initial 10 to 60, 50, 40:** Reducing analyses to subjects with initial FM scores of 10 – 60 only marginally improved parameter estimation performance, i.e., up to 5 points *constant* recovery. Considering subjects with initial FM scores of 10 – 50 and 10 – 40 allowed accurate estimation up to 10 – 15 points and 25 points, respectively. Noteworthily, estimation performance dropped once the percentage of subjects at ceiling exceeded 7 %.

Exclusively looking at subjects in the range of *FM-initial* 10 – 60 reduced the empirical number of subjects at absolute ceiling from 37 to 18 (15.2 %and 9.0 %, respectively), in the range of *FM-initial* 10 – 50 to 7 (4.7 %) and in the range of *FM-initial* 10 – 40 to 4 (4.3 %).

**Supplementary table 3a: Constant recovery: No ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 4.99 | 0.70 | 0.95 | 1.00 | 1.00 | 0.69 | 0.95 |
| 10 | 10.00 | 10.00 | 0.67 | 0.94 | 1.00 | 1.00 | 0.68 | 0.95 |
| 15 | 15.00 | 14.97 | 0.70 | 0.94 | 1.00 | 1.00 | 0.68 | 0.95 |
| 20 | 20.00 | 19.99 | 0.68 | 0.94 | 1.00 | 1.00 | 0.67 | 0.94 |
| 25 | 25.00 | 24.99 | 0.71 | 0.95 | 1.00 | 1.00 | 0.70 | 0.95 |
| 30 | 30.00 | 29.98 | 0.69 | 0.95 | 1.00 | 1.00 | 0.67 | 0.95 |
| 35 | 35.00 | 34.99 | 0.69 | 0.95 | 1.00 | 1.00 | 0.68 | 0.95 |
| 40 | 40.00 | 40.01 | 0.69 | 0.96 | 1.00 | 1.00 | 0.66 | 0.95 |
| 45 | 45.00 | 45.03 | 0.70 | 0.95 | 1.00 | 1.00 | 0.68 | 0.95 |
| 50 | 50.00 | 50.01 | 0.69 | 0.95 | 1.00 | 1.00 | 0.68 | 0.95 |

|  |  |  |
| --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |
| 1.04 | 0.96 | 0.00 |

**Supplementary table 3b: Constant recovery: Ceiling enforced.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 6.37 | 0.25 | 0.60 | 1.00 | 0.95 | 0.01 | 0.17 |
| 10 | 10.00 | 13.15 | 0.00 | 0.03 | 1.00 | 0.88 | 0.00 | 0.00 |
| 15 | 15.00 | 20.43 | 0.00 | 0.00 | 1.00 | 0.78 | 0.00 | 0.00 |
| 20 | 20.00 | 27.88 | 0.00 | 0.00 | 1.00 | 0.67 | 0.00 | 0.00 |
| 25 | 25.00 | 35.35 | 0.00 | 0.00 | 1.00 | 0.55 | 0.00 | 0.00 |
| 30 | 30.00 | 42.46 | 0.00 | 0.00 | 1.00 | 0.42 | 0.00 | 0.00 |
| 35 | 35.00 | 49.05 | 0.00 | 0.00 | 1.00 | 0.30 | 0.00 | 0.00 |
| 40 | 40.00 | 54.85 | 0.00 | 0.00 | 1.00 | 0.20 | 0.00 | 0.00 |
| 45 | 45.00 | 59.56 | 0.00 | 0.00 | 1.00 | 0.11 | 0.00 | 0.00 |
| 50 | 50.00 | 62.81 | 0.00 | 0.00 | 1.00 | 0.05 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 0.99 | 0.96 | -0.18 | 20 |
| 0.92 | 0.95 | -0.40 | 34 |
| 0.83 | 0.94 | -0.60 | 46 |
| 0.73 | 0.91 | -0.74 | 57 |
| 0.62 | 0.88 | -0.84 | 66 |
| 0.51 | 0.83 | -0.90 | 74 |
| 0.40 | 0.76 | -0.94 | 82 |
| 0.29 | 0.67 | -0.96 | 90 |
| 0.20 | 0.54 | -0.98 | 99 |
| 0.13 | 0.24 | -0.99 | 110 |

**Supplementary table 3c: Constant recovery: Ceiling enforced, Subset 10 - 60.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 5.39 | 0.64 | 0.93 | 1.00 | 0.98 | 0.57 | 0.93 |
| 10 | 10.00 | 11.36 | 0.28 | 0.65 | 1.00 | 0.94 | 0.04 | 0.27 |
| 15 | 15.00 | 18.08 | 0.01 | 0.06 | 1.00 | 0.87 | 0.00 | 0.00 |
| 20 | 20.00 | 25.36 | 0.00 | 0.00 | 1.00 | 0.76 | 0.00 | 0.00 |
| 25 | 25.00 | 32.93 | 0.00 | 0.00 | 1.00 | 0.64 | 0.00 | 0.00 |
| 30 | 30.00 | 40.39 | 0.00 | 0.00 | 1.00 | 0.50 | 0.00 | 0.00 |
| 35 | 35.00 | 47.44 | 0.00 | 0.00 | 1.00 | 0.36 | 0.00 | 0.00 |
| 40 | 40.00 | 53.69 | 0.00 | 0.00 | 1.00 | 0.24 | 0.00 | 0.00 |
| 45 | 45.00 | 58.94 | 0.00 | 0.00 | 1.00 | 0.13 | 0.00 | 0.00 |
| 50 | 50.00 | 62.53 | 0.00 | 0.00 | 1.00 | 0.06 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.04 | 0.95 | -0.05 | 7 |
| 1.00 | 0.95 | -0.17 | 16 |
| 0.92 | 0.94 | -0.39 | 28 |
| 0.82 | 0.92 | -0.60 | 39 |
| 0.71 | 0.89 | -0.75 | 48 |
| 0.59 | 0.85 | -0.85 | 55 |
| 0.46 | 0.78 | -0.91 | 63 |
| 0.35 | 0.69 | -0.95 | 72 |
| 0.24 | 0.55 | -0.97 | 81 |
| 0.15 | 0.38 | -0.99 | 91 |

**Supplementary table 3d: Constant recovery: Ceiling enforced, Subset 10 - 50.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 5.02 | 0.67 | 0.96 | 1.00 | 1.00 | 0.69 | 0.95 |
| 10 | 10.00 | 10.12 | 0.68 | 0.95 | 1.00 | 0.99 | 0.69 | 0.96 |
| 15 | 15.00 | 15.46 | 0.66 | 0.94 | 1.00 | 0.98 | 0.61 | 0.91 |
| 20 | 20.00 | 21.63 | 0.28 | 0.64 | 1.00 | 0.92 | 0.05 | 0.26 |
| 25 | 25.00 | 28.71 | 0.00 | 0.05 | 1.00 | 0.81 | 0.00 | 0.00 |
| 30 | 30.00 | 36.35 | 0.00 | 0.00 | 1.00 | 0.67 | 0.00 | 0.00 |
| 35 | 35.00 | 44.00 | 0.00 | 0.00 | 1.00 | 0.51 | 0.00 | 0.00 |
| 40 | 40.00 | 51.23 | 0.00 | 0.00 | 1.00 | 0.34 | 0.00 | 0.00 |
| 45 | 45.00 | 57.42 | 0.00 | 0.00 | 1.00 | 0.20 | 0.00 | 0.00 |
| **50** | 50.00 | 61.87 | 0.00 | 0.00 | 1.00 | 0.09 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.09 | 0.92 | 0.00 | 0 |
| 1.08 | 0.92 | -0.01 | 1 |
| 1.06 | 0.92 | -0.05 | 6 |
| 1.00 | 0.92 | -0.20 | 13 |
| 0.90 | 0.90 | -0.43 | 21 |
| 0.77 | 0.87 | -0.65 | 29 |
| 0.63 | 0.81 | -0.80 | 37 |
| 0.48 | 0.72 | -0.89 | 45 |
| 0.33 | 0.58 | -0.95 | 55 |
| 0.21 | 0.40 | -0.98 | 64 |

**Supplementary table 3e: Constant recovery: Ceiling enforced, Subset 10 - 40.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 5.00 | 0.69 | 0.94 | 1.00 | 1.00 | 0.68 | 0.95 |
| 10 | 10.00 | 9.99 | 0.68 | 0.95 | 1.00 | 1.00 | 0.66 | 0.96 |
| 15 | 15.00 | 14.92 | 0.70 | 0.96 | 1.00 | 1.00 | 0.69 | 0.95 |
| 20 | 20.00 | 20.11 | 0.70 | 0.95 | 1.00 | 0.99 | 0.71 | 0.96 |
| 25 | 25.00 | 25.47 | 0.66 | 0.94 | 1.00 | 0.97 | 0.62 | 0.93 |
| 30 | 30.00 | 31.69 | 0.41 | 0.79 | 1.00 | 0.90 | 0.19 | 0.60 |
| 35 | 35.00 | 39.14 | 0.02 | 0.13 | 1.00 | 0.75 | 0.00 | 0.00 |
| 40 | 40.00 | 47.22 | 0.00 | 0.00 | 1.00 | 0.54 | 0.00 | 0.00 |
| 45 | 45.00 | 54.82 | 0.00 | 0.00 | 1.00 | 0.32 | 0.00 | 0.00 |
| 50 | 50.00 | 60.58 | 0.00 | 0.00 | 1.00 | 0.15 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.16 | 0.86 | 0.00 | 0.00 |
| 1.16 | 0.86 | 0.00 | 0.00 |
| 1.17 | 0.86 | 0.01 | 0.00 |
| 1.16 | 0.86 | -0.01 | 1.00 |
| 1.13 | 0.86 | -0.05 | 3.00 |
| 1.05 | 0.86 | -0.18 | 8.00 |
| 0.91 | 0.82 | -0.44 | 14.00 |
| 0.72 | 0.75 | -0.69 | 20.00 |
| 0.52 | 0.62 | -0.86 | 28.00 |
| 0.33 | 0.44 | -0.94 | 36.00 |

**Supplementary table 3f: Constant recovery: Ceiling enforced, Subset 10 - 45.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Amount of recovery** | **True Intercept** | **Estimated intercept** | **Met 68%-CI** | **Met 95%-CI** | **True Slope** | **Estimated slope** | **Met 68%-CI** | **Met 95%-CI** |
| 5 | 5.00 | 4.94 | 0.69 | 0.94 | 1.00 | 1.00 | 0.68 | 0.94 |
| 10 | 10.00 | 9.97 | 0.68 | 0.95 | 1.00 | 1.00 | 0.69 | 0.95 |
| 15 | 15.00 | 15.13 | 0.69 | 0.95 | 1.00 | 0.99 | 0.71 | 0.96 |
| 20 | 20.00 | 20.56 | 0.65 | 0.95 | 1.00 | 0.97 | 0.62 | 0.93 |
| 25 | 25.00 | 26.73 | 0.31 | 0.68 | 1.00 | 0.91 | 0.09 | 0.45 |
| 30 | 30.00 | 33.81 | 0.01 | 0.11 | 1.00 | 0.79 | 0.00 | 0.00 |
| 35 | 35.00 | 41.46 | 0.00 | 0.00 | 1.00 | 0.63 | 0.00 | 0.00 |
| 40 | 40.00 | 49.27 | 0.00 | 0.00 | 1.00 | 0.44 | 0.00 | 0.00 |
| 45 | 45.00 | 56.05 | 0.00 | 0.00 | 1.00 | 0.26 | 0.00 | 0.00 |
| 50 | 50.00 | 61.18 | 0.00 | 0.00 | 1.00 | 0.12 | 0.00 | 0.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Variability ratio** | **Correlation X - Y** | **Correlation X - Change** | **Numbers of subjects at or above absolute ceiling** |
| 1.12 | 0.89 | 0.00 | 0.00 |
| 1.12 | 0.89 | 0.00 | 0.00 |
| 1.11 | 0.89 | -0.01 | 1.00 |
| 1.09 | 0.89 | -0.06 | 4.00 |
| 1.02 | 0.89 | -0.19 | 9.00 |
| 0.91 | 0.87 | -0.42 | 15.00 |
| 0.77 | 0.82 | -0.65 | 23.00 |
| 0.59 | 0.74 | -0.81 | 30.00 |
| 0.42 | 0.61 | -0.91 | 39.00 |
| 0.27 | 0.43 | -0.96 | 47.00 |

**8 Bayesian hierarchical models *FM-initial* 10 – 45 for Fitters: Model comparison based on the WAIC and interpretation of marginal posterior distributions (c.f. section Results in Main Body)**

**A screenshot of a cell phone

Description automatically generated**

**Supplementary Figure 3.** Subset *FM-Initial* 10 – 45 for *fitters*. Model comparison based on the widely applicable criterion (WAIC), resulting in a similar model ranking in comparison to the leave-one-out-cross-validations (LOOCV) results. The standard-form regression model is top-ranked, closely followed by the *proportional to lost function* model and then *constant* recovery model. The error bars indicate associated standard errors with WAIC-estimates, in-sample estimates as well as the estimates of differences between two WAIC-estimates

**Supplementary Figure OOOO. Subset FM-Initial 10 – 45. A** Marginal posteriors for slope parameters in the proportional to lost function model **B** Difference of extreme individual studies’ parameter sets on the lower hierarchy level. Given the dispersion of individual studies’ marginal posterior’s slope parameters in A, we here visually compare Study 6’s (Guggisberg et al.) and Study 4’s sampled posteriors. As the zero (which would be the mean of the difference between similar marginal posteriors) is at the outer tail of the distribution of difference, the difference between the two marginal slope posteriors is substantial.

**9 Further Discussion of Fitting Results (c.f. section Results in Main Body)**

Here, we provide further interpretation of the fitting results that we obtain for our human data. There are three cases where our fitting gives parameter estimates that are just outside the range of accurate retrieval according to our ground-truth (synthetic data) simulations. This happens for change-form *proportional to spared* for both *fitters* alone and *fitters & non-fitters* combined, and change-form *constant* recovery for *fitters* alone. Importantly, these are all losing models. That is, for *fitters*, change-form *proportional to spared* was the worst fitting followed by change-form c*onstant* recovery, while for *fitters & non-fitters*, change-form *proportional to spared* was the worst model.

Since these are losing models, the key inferential error that we want to defend ourselves against is that we have *under*-estimated the R-squared and model evidence for these, since only this would alter our conclusions, as to which are the winning models. Critically, though, in all these cases, any inaccuracy in fitting would work in the opposite direction, i.e., to *over*-estimate the quality of fit. This is very simple because if the functional-form of the model being fit was that of the true generating model, the correct parameter setting would be in the parameter space being searched/fit. The parameter values selected in the model fitting have to be the correct settings or “better” fittings than them, otherwise, they would not have been selected over and above the true ones by the fitting algorithm.

Thus, in all these critical cases, the R-squared and model evidence can only be correct or *over*-estimated, *under*-estimation is not possible. However, even if there is *over*-estimation, the models still lose, confirming that the inferences we have made are sound. We discuss these three cases in a little more depth next.

*Constant Recovery*: when we fit the *constant* recovery rule in its change formulation (see main-body **Figure 2[C]**) to the *fitters* group, we obtain an R-squared of 5.8% and a constant of 26pt (see main-body section “Final Model Comparison (on human data) in the Subset of FM-initial 10-45”). In supplementary material section 7 (“Detailed Evaluation of Simulations”), we show synthetic data simulation results for the different models. The simulations for *constant* recovery, see **Supplementary Table 3f**, show that with 10-45 subsetting, we get very accurate parameter retrieval for constants up to 20 points. For a constant of 25, though, retrieval becomes less accurate and worsens further for 30 and beyond.

The central inference that we are seeking to make in our analysis of fitters is that standard-form regression has the highest R-squared and model evidence, followed by *proportional to lost* function and then *constant* recovery. Thus, as just discussed, the scenario that we wish to defend ourselves against is that we have underestimated the R-squared and model evidence for the fit of *constant* recovery. For the reasons just discussed, we can argue that such an underestimation has not been made. This, then, suggests that our conclusion that standard-form and *proportional to lost* fit better than *constant* recovery for the 10-45 subset of the *fitters* data, is indeed a reliable inference.

For the *fitters & non-fitters* combined data set, *constant* recovery obtains a constant of ~20points, putting it in the range of reliable parameter retrieval.

*Proportional to Spared*: When we fit the *proportional to spared* model in the change formulation (see main-body **Figure 2[B]**) for both *fitters* alone and *fitters & non-fitters*, we obtain slopes that are in the upper part of the allowable range, where retrievability of parameters starts to breakdown; see **Supplementary Table 2f,** for 10-45 subsetting, which is around 60%. However, both for *fitters* and *fitters & non-fitters*, when this arises, we can use the argument just made to justify that an under-estimate of fit has not been made, supporting our inferences that models other than *proportional to spared* are the best fitting.

**10 Additional subset evaluations for Fitters: Bayesian hierarchical models *FM-initial* 10 – 40 and *FM-initial* 10 – 50 (c.f. Results section of Main BodyA close up of a map

Description automatically generated**

**Supplementary Figure 4. Alleviating confounds by ceiling effects and mathematical coupling: Marginal posteriors for Fitters of the hierarchical models’ parameters considering all subjects with initial FM scores between ten and 40. Final model comparison.** **[A]** *Proportional to lost* function recovery (formulated as in main-body **Figure 2[A])**; **[B]** *Proportional to spared* function recovery (formulated as in main-body **Figure 2[B]**); **[C]** *Constant* recovery (formulated as in main-body **Figure 2[C]**); **[D]** (Unconstrained) Standard-form regression**; [E]** Final Bayesian model comparison relying on leave-one-out-cross-validation (LOOCV). Empty circles represent the LOOCV-score, black error bars the corresponding standard error. The filled black circles mark the models’ in-sample deviance, grey triangles the difference to the top-ranked model, as well as the associated standard error. Lastly, the lowest LOOCV-value is indicated by the vertical dashed grey line. The standard-form regression model provides the best out-of-sample performance and is ranked first in the model comparison, closely followed by the *constant* and *proportional to lost function* recovery models.

A close up of a map

Description automatically generated

**Supplementary Figure 5. Alleviating confounds by ceiling effects and mathematical coupling: Marginal posteriors for Fitters of the hierarchical models’ parameters considering all subjects with initial FM scores between ten and 50. Final model comparison. [A]** *Proportional to lost function* recovery (formulated as in main-body **Figure 2[A]**); **[B]** *Proportional to spared function* recovery (formulated as in main-body **Figure 2[B]**); **[C]** *Constant* recovery (formulated as in main-body **Figure 2[C]**); **[D]** (Unconstrained) Standard-form regression; **[E]** Final Bayesian model comparison relying on leave-one-out-cross-validation (LOOCV). Empty circles represent the LOOCV-Score, black error bars the corresponding standard error. The filled black circles mark the models’ in-sample deviance, grey triangles the difference to the top-ranked model, as well as the associated standard error. Lastly, the lowest LOOCV-value is indicated by the vertical dashed grey line. The standard-form regression model provides the best out-of-sample performance and is ranked first in the model comparison, immediately followed by the *proportional to lost function* recovery model. The *constant* recovery model follows with quite some distance, the *proportional to spared* recovery model is ranked last, as in the other subsets’ model comparisons.

Gladstone, D. J., Danells, C. J., & Black, S. E. (2002). The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabilitation and Neural Repair*, *16*(3), 232–240.

Guggisberg, A. G., Nicolo, P., Cohen, L. G., Schnider, A., & Buch, E. R. (2017). Longitudinal structural and functional differences between proportional and poor motor recovery after stroke. *Neurorehabilitation and Neural Repair*, *31*(12), 1029–1041.

Hawe, R. L., Scott, S. H., & Dukelow, S. P. (2019). Taking Proportional Out of Stroke Recovery. *Stroke*, *50*(1), 204–211.

Prabhakaran, S., Zarahn, E., Riley, C., Speizer, A., Chong, J. Y., Lazar, R. M., … Krakauer, J. W. (2008). Inter-individual Variability in the Capacity for Motor Recovery After Ischemic Stroke. *Neurorehabilitation and Neural Repair*, *22*(1), 64–71. https://doi.org/10.1177/1545968307305302

Zarahn, E., Alon, L., Ryan, S. L., Lazar, R. M., Vry, M.-S., Weiller, C., … Krakauer, J. W. (2011). Prediction of Motor Recovery Using Initial Impairment and fMRI 48 h Poststroke. *Cerebral Cortex*, *21*(12), 2712–2721. https://doi.org/10.1093/cercor/bhr047