SUPPLEMENTAL INFORMATION

Altered Global Brain Signal in Schizophrenia

Genevieve J. Yang[1,2,3] *, John D. Murray[4] *, Grega Repovs[5], Michael W. Cole[6], Aleksandar Savic[1,3,7], Matthew F. Glasser[12], Christopher Pittenger[1,2,3,11], John H. Krystal[1,3,8], Xiao-Jing, Wang[4,13], Godfrey D. Pearlson[1,9,10], David C. Glahn[1,10], Alan Anticevic[1,2,3,8,11]

1. Department of Psychiatry, Yale University School of Medicine, 300 George Street, New Haven, CT 06511, USA
2. Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06520, USA
3. Abraham Ribicoff Research Facilities, Connecticut Mental Health Center, New Haven, CT 06519, USA
4. Center for Neural Science, New York University, New York, NY 06510, USA
5. Department of Psychology, University of Ljubljana, Ljubljana, Slovenia
6. Center for Molecular & Behavioral Neuroscience, Rutgers University, Newark, NJ 07102, USA
7. University Psychiatric Hospital Vrapce, University of Zagreb, Zagreb 10000, Croatia
8. NIAAA Center for the Translational Neuroscience of Alcoholism, New Haven, CT 06519, USA
9. Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA
10. Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, 200 Retreat Avenue, CT 06106, USA
11. Department of Psychology, Yale University, 2 Hillhouse Avenue, CT 06520, USA
12. Department of Anatomy and Neurobiology, Washington University in St. Louis, MO 63130, USA
13. NYU-ECNU Joint Institute of Brain and Cognitive Science, NYU-Shanghai, Shanghai, China

* GJY and JDM contributed equally to this manuscript

Corresponding Author:

Alan Anticevic Ph.D., Yale University, Department of Psychiatry
34 Park St., New Haven, CT 06519, Office (203) 974-7763
alan.anticevic@yale.edu

Running Title: Global Signal in Schizophrenia

INVENTORY OF SUPPLEMENTAL INFORMATION [SI]

SI EXPERIMENTAL PROCEDURES

Participant Inclusion Criteria and Characteristics. Here we studied three independent clinical samples: i) 90 chronic SCZ patients and 90 demographically matched HCS (**Table S1**), to 'discover' all effects. ii) replication sample consisting of 71 SCZ patients and 74 demographically matched HCS obtained from a publicly-distributed dataset provided by the Center for Biomedical Research Excellence (COBRE) (http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html) (**Table S2**); iii) 73 patients diagnosed with BD and 56 HCS, to test the diagnostic specificity of SCZ effects(6) (**Table S3**). Across samples, all subjects met identical neuroimaging exclusion criteria, underwent identical preprocessing and analyses (see analysis sections below). The 'discovery' sample schizophrenia (SCZ) participants (N=90) were identified through outpatient clinics and community mental health facilities in the Hartford (CT) area. Complete recruitment details for this sample are described in our prior work(1). Briefly, patient inclusion criteria were as follows: i) SCZ diagnosis as determined by the Structured Clinical Interview (SCID) for the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)(2), administered by experienced MA or PhD-level research clinicians; ii) no major medical or neurological conditions (e.g. epilepsy, migraine, head trauma with loss of consciousness); and iii) IQ>70 assessed by widely-accepted methods for estimating premorbid intelligence levels [either National Adult Reading Test (NART), Wide Range Achievement Test (WRAT) or Wechsler Test of Adult Reading (WTAR) depending on the study protocol](3). As in our prior studies, these measures were normed and converted to IQ equivalents for each subject. If more than one premorbid achievement measure was available per subject the scaled scores were averaged per standard practice(4). An important consideration in the present study is whether identified effects in any way relate to cognitive impairment in SCZ. Related to this question, the 'IQ' measures are based on measures that assess premorbid academic achievement levels and do not capture the more complex higher-order cognitive deficits(5). While this question is beyond the scope of our investigation, future studies should systematically examine if cognitive impairment relates to presently identified effects, In the present study we did not exclude patients with a lifetime co-morbid Axis I anxiety disorders and/or history of substance abuse in the schizophrenia sample to ensure an inclusive and representative sample of patients(6). However, all discovery SCZ sample participants were required to be fully remitted >6 months prior to the study.

Healthy comparison subjects (HCS) (N=90) were recruited through media advertisements and flyers posted in the Medical Center area. Inclusion criteria for HCS were: i) no current or lifetime Axis I psychiatric disorder as assessed by SCID-NP; ii) no history of medical or neurological conditions; and iii) no history of psychotic disorders in first-degree relatives (reported by detailed family history). The bipolar disorder (BD) sample participants (N=73) were recruited at the same site as the SCZ discovery sample. The same general exclusion/inclusion criteria were applied, described previously in more detail(7-9).

As noted, the SCZ 'replication' sample was provided to the neuroimaging community by COBRE researchers. Critically, this large and independent SCZ sample has been extensively characterized, demographically matched and quality-assured across a number of prior reports (http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html). Replication sample SCZ patients were excluded if they had: i) history of neurological disorder, ii) history of mental retardation, ii) history of severe head trauma with more than 5 minutes loss of consciousness; iv) history of substance abuse or dependence within the last 12 months. Diagnostic decisions were reached using the SCID interview for the DSM-IV. Collectively, these criteria and demographic are highly comparable across the two SCZ samples.

To allow full inspection of these clinical samples we provide complete demographics details in **Tables S1-3**. Across all samples we accomplished matching on a number of relevant demographic variables, ensuring comparable between-group demographics. The discovery sample groups did not significantly differ on any of the variables (see **Table S1**), apart from educational attainment and premorbid intellectual functioning (IQ) (see aforementioned comments). IQ was lower for SCZ patients, and likely reflects the shortened educational achievement for patients due to illness onset(10). Also, education differences are impacted by the illness course and thus were not considered as a covariate as they likely reflect illness trajectory. Notably, alcohol/drug use, age, gender, smoking status and medication did not alter reported effects for the discovery sample (see **Fig. S3** for comprehensive confound analyses). 75/90 schizophrenia patients were receiving antipsychotic treatment. SCZ patients for the COBRE replication sample were also receiving stable doses of antipsychotic medication with no medication changes for at least 1 month. All medication were converted to

chlorpromazine (CPZ) equivalents(11) and verified by trained raters (AA & AS). We used CPZ levels as a co-variate for our discovery analyses and we verified that the CPZ values did not correlate with the CGm power and variance main effects (see **Fig. S3**). Critically, the two SCZ samples that were recruited across different sites received somewhat different levels of medication doses. Yet, the key effects replicated across both SCZ patient samples. Collectively, these results are inconsistent with medication being a major confound in our analyses.

Schizophrenia Symptoms. Schizophrenia symptom severity, across both samples, was determined using the Positive and Negative Syndrome Scale (PANSS), a widely-used symptom instrument, which captures positive, negative and general psychopathology symptom dimensions(12) (**Table S1**). BD patients were in remission at the time of the scan(7), therefore no symptom analyses were attempted.

Neuroimaging Data Acquisition. The SCZ discovery, the BD sample and the their respective HCS participated in neuroimaging data collection at the Olin Neuropsychiatry Research Center using a Siemens-Allegra 3T scanner with identical acquisition parameters(7). Images sensitive to blood oxygenation level-dependent (BOLD) signal were acquired with axial slices parallel to the anterior-posterior commissure (AC-PC) using a T2*-weighted gradient-echo, echo-planar sequence (TR/TE=1500/27ms, flip angle=60°, field of view=24x24 cm, acquisition matrix=64x64, voxel size=3.43x3.43x4mm) covering the whole brain. The acquisition lasted 5.25 minutes and produced 210 volumetric images per subject (29 slices/volume, inter-slice gap=1mm). Subjects were instructed to lay awake in the scanner and keep their eyes open. Subjects were monitored on a video camera to ensure that they stayed awake and were removed from the analyses if they fell asleep during the scan or if their head movement exceeded 1mm along any axis. Structural images were acquired using a T1-weighted, 3D magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR/TE/TI=2200/4.13/766ms, flip angle=13°, voxel size [isotropic]=.8mm, image size=240x320x208 voxels), with axial slices parallel to the AC-PC line. SCZ patients comprising the replication sample (N=71) and their respective HCS (N=74) underwent data collection at Center for Biomedical Research Excellence using a Siemens Tim-Trio 3T scanner. Full acquisition details for the SCZ replication sample and HCS is detailed previously(13-15). Briefly, BOLD signal was collected with 32 axial slices parallel to the AC-PC using a T2*-weighted gradient-echo, echo-planar sequence (TR/TE=2000/29ms, flip angle=75°, acquisition matrix=64x64, voxel size=3x3x4mm). The acquisition lasted 5 minutes and produced 150 volumetric images per subject. Structural images were acquired using a 6 minute T1-weighted, 3D MPRAGE sequence (TR/TE/TI=2530/[1.64, 3.5, 5.36, 7.22, 9.08]/900, flip angle=7°, voxel size [isotropic]=1mm, image size=256x256x176 voxels), with axial slices parallel to the AC-PC line. All the described parameters were provided via the publically-distributed website (http://fcon_1000..projects.nitrc.org/indi/retro/cobre.html).

Neuroimaging Preprocessing & Analysis. Preprocessing followed prior validated approaches that were applied to clinical populations(7, 9). Critically, we performed identical preprocessing procedures across all subjects collected across scanners, which were then registered to the same common standard space and interpolated to the same resolution. We first performed the following preprocessing steps for all BOLD images, as done in our prior studies(7, 9): i) slice-time correction, ii) first 5 images removed from each run, iii) rigid body motion correction, iv) 12-parameter affine transform of the structural image to the Talairach coordinate system, and v) co-registration of volumes to the structural image with 3x3x3mm re-sampling, ensuring all BOLD images across both scanners were interpolated to the same resolution.

In addition, all BOLD images had to pass stringent quality assurance criteria to ensure that all functional data were of comparable and high quality: i) signal-to-noise ratios (SNR)>100, computed by obtaining the mean signal and standard deviation (sd) for a given slice across the BOLD run, while excluding all non-brain voxels across all frames(7). The final samples had high SNR values (see **Tables S1-3**); ii) movement scrubbing as recommended by Power et al.(16, 17). As accomplished previously(18), all image frames with possible movement-induced artifactual fluctuations in intensity were identified via two criteria: First, frames in which sum of the displacement across all 6 rigid body movement correction parameters exceeded 0.5mm (assuming 50mm cortical sphere radius) were identified; Second, root mean square (RMS) of differences in intensity between the current and preceding frame was computed across all voxels divided by mean intensity

and normalized to time series median. Frames in which normalized RMS exceeded the value of 3 were identified. The frames flagged by either criterion were marked for exclusion (logical or), as well as the one preceding and two frames following the flagged frame. Movement scrubbing was performed for all reported analyses across all subjects except for the power frequency analyses. The scrubbing step was omitted in this case, because removing arbitrary frames would have corrupted the time-dependent frequency information of the signal. Nevertheless, all effects remained robust with complementary analyses techniques (i.e. variance analyses) following stringent movement scrubbing and after explicitly controlling for movement (see **Fig. S3**). Collectively, these quality assurances add confidence that typical neuroimaging confounds (i.e. SNR or movement) are not driving present effect.

Lastly, to remove spurious signal in resting-state data we completed additional preprocessing steps, as is standard practice(19): all BOLD time-series underwent high (>0.009 Hz) and low (<0.08 Hz) pass temporal filtering, removal of nuisance signal extracted from anatomically-defined ventricles, white matter, and the remaining brain voxels (i.e. global signal) (all identified via individual-specific FreeSurfer segmentations(20)), as well as 6 rigid-body motion correction parameters, and their first derivatives using previously validated in-house Matlab tools(21).

Where appropriate, type I error correction was implemented via threshold-free cluster enhancement (TFCE) non-parametric techniques implemented in FSL's *Randomise* tool(22) (e.g. for the GBC analyses, see below for details). Results were visualized using Caret 5.5 (http://brainvis.wustl.edu/wiki/index.php/Caret) and NeuroLens software (http://www.neurolens.org).

Global Signal Regression (GSR). GSR was performed using standard procedures(23) by calculating mean raw BOLD signal averaged over all voxels for each time point, explicitly excluding ventricles and white matter signal (which are defined as separate nuisance regressors). The GS and its first derivative (with respect to time) were used as nuisance predictor terms within a multiple linear regression model along with other nuisance predictor terms (ventricular signal, white matter signal, movement parameters, and the first derivatives of each of these, as noted above).

Power Analysis. To quantify the impact of GSR on the frequency content of CGm and Gm BOLD signal, we evaluated frequency amplitudes using the multi-taper periodogram method, available as part of the Nitime Time-Series Analysis library (http://nipy.org/nitime/)(24). The first 5 volumes were discarded to allow for stable BOLD signal levels. We quantified the CGm power in each group and preprocessing condition by computing each individual's average CGm frequency amplitude across all frequencies recorded from the periodogram of the CGm (or Gm) signal. The resulting average CGm power calculated for each individual was entered as a dependent measure into all 2nd level ANOVA analyses.

CGm/Gm Signal and Thalamic Isolation. Across a number of analyses we quantified the alterations in CGm/Gm signal in SCZ versus BD as a function of GSR. Moreover, we computed the relationship between the CGm and anatomically-defined thalamic nuclei for each individual (see below). To extract appropriate Gm/CGm or thalamic signal for each individual we used individual-specific FreeSurfer segmentations(20), which were visually inspected for quality by a trained rater (AA) (**Table S4** lists all FreeSurfer codes used across analyses). We extracted the BOLD signal from regions labeled as belonging to CGm/Gm for each individual, which was then used in all subsequent analyses. Similarly, we identified thalamic BOLD signal by extracting average time-series across all voxels in each subject's bilateral thalamus, as done previously(1).

Whole-brain Voxel-wise Variance Analysis. To test the hypothesis that voxel-wise whole-brain variance may be altered in SCZ, we computed the variability of each voxel's time-series with and without GSR. This analysis is distinct from the variability of the CGm in that it computes the variance of each voxel, as opposed to the, spatial average across all Gm/CGm voxels. We computed the variability of each voxel's time series after all noted preprocessing steps, including movement scrubbing. We did so for all SCZ patients, combining the discovery and replication samples (given a highly similar pattern of results) (N=161). We repeated the voxel-wise analyses for BD patients to establish diagnostic specificity. Following the variance calculation, we computed a whole-brain voxel-wise between-group t-test for SCZ patients relative to HCS, treating each

voxel's variance as the dependent variable. That is, we obtained a voxel-wise variance map for each subject, where each voxel's value represents its variability over time.

PFC Connectivity Analysis. The data-driven *global brain connectivity* (GBC) approach(9, 25) was applied using in-house Matlab tools, here restricted to the PFC (referred to hereafter as *restricted global brain connectivity* - rGBC). We focused on the PFC given numerous prior studies implicating PFC dysfunction in SCZ(26), and to remain maximally powered given the smaller search space of PFC voxels. Furthermore, our prior reports identified PFC rGBC alterations in chronic SCZ, which served as a basis for comparison. Importantly, the PFC rGBC analysis is fully data-driven, adding a complementary technique to the thalamic seed-driven analyses. Specifically, rGBC provides a sensitive measure by computing correlational strength of each voxel with all other voxels being considered. As previously described(7, 9, 25), GBC is designed to estimate the connectivity between each individual voxel and every other voxel in the brain or a restricted search space. Thus, rGBC is optimized to estimate connectivity at every voxel with every other voxel in a specific search space (in this case PFC).

As done previously(7, 9), we conducted rGBC analysis restricted to voxels only within subject-specific anatomically defined PFC gray matter masks defined via Freesurfer software(20) (see **Table S4** for a list of all relevant FreeSurfer codes). All PFC Freesurfer segmentations were also visually inspected for quality by a trained rater (AA). Before rGBC analysis, BOLD signal within the subject-specific cortical mask was spatially smoothed with a 6mm full-width-at-half-maximum (FWHM) Gaussian kernel and dilated by two voxels (6mm) to account for individual differences in anatomy. Finally, for each PFC voxel, we computed a correlation with every other PFC voxel, transformed the correlations to Fisher z-values, and finally computed their mean. This calculation yielded an rGBC map for each subject where each voxel value represents the mean connectivity of that voxel with the rest of PFC (see **Fig. 4**).

Quantification of Between-Group Difference Map Similarity and Overlap. The data-driven PFC rGBC analyses (**Fig. 5**) suggest that GSR may affect the pattern of between-group differences in complex ways that result in distinct between-group effects before and after this processing step is applied. Because of the larger magnitude of the GS variability in SCZ, this could lead to two distinct possibilities, as noted in the main text: i) GSR could have introduced a 'uniform' shift in the data that is just larger in magnitude for the SCZ group; ii) GSR could have introduced a 'non-uniform' shift such that the effect is differentially distributed across voxels for one group versus the other, a possibility articulated by mathematical simulations(27). We designed two analyses to arbitrate between these possibilities:

1) Testing for Similarity of Between-Group rGBC Maps Before and After GSR: If the spatial configuration of the rGBC is non-uniformly affected by GSR then the between-group maps should be qualitatively distinct (i.e. they should not be similar before and after GSR). Conversely, if the pattern is largely similar but just uniformly shifted by a rigid transform (see **Fig. 4E**), then the statistical similarity of the maps should be high (without any statistical threshold employed). To distinguish between these hypotheses and assess the similarity of between-group difference maps pre/post GSR via a single value, we made use of the eta^2 index that quantifies the pattern of similarity between two signals, or in this case functional images. Briefly eta^2 is designed to vary from 0 for no similarity to 1 for perfect signal similarity; for further detail see(28). That is, the employed method was explicitly designed to capture similarity (eta^2) between a pair of functional connectivity maps(28). To accomplish this calculation we converted the Fz between-group difference maps to one-dimensional vectors for each PFC voxel. Next, we computed eta^2 between vectors before and after GSR. This calculation captured the similarity across the two between-group maps before and after GSR without any statistical threshold employed. In addition to eta^2 we also used standard r-values. However, there are important benefits of using eta^2: whereas r reflects only the dissimilarities in the pattern of connectivity, eta^2 also captures possible differences in the connectivity strength for voxels with similar connectivity patterns. Therefore, the resulting eta^2 captures the overall 'similarity' between two connectivity maps – in this case the maps showing between-group differences before and after GSR:

$$1 - eta^2 = 1 - (1 - \frac{\sum_{i=1}^{n}[(a_i - m_i)^2 + (b_i - m_i)^2]}{\sum_{i=1}^{n}[(a_i - \bar{M}_i)^2 + (b_i - \bar{M}_i)^2]}$$

As described in detail by Cohen and colleagues(28), here *ai and bi* correspond to connectivity at position *i* (in this case a given voxel in the PFC) for maps *a* and *b* respectively (in this case the between-group difference for patients vs. control before (map A) and after (map B) GSR). *mi* corresponds to the mean value of the two images at position *i*, (*ai*+*bi*)/2. M-bar represents the grand mean value across the mean image (designated by *m*). By computing, eta^2 and r-values we are able to quantitatively assess the degree of similarity between resulting maps pre and post GSR (see **Fig. S8A-B**).

2) Testing for Percentage of Overlapping Voxels of Between-Group rGBC Maps Before and After GSR: The eta^2 and r-value calculations above provide a single index of statistical similarity before and after GSR. However, to further assess if the similarity is dominated by a certain 'range' of values in the statistical map, we quantified the percent of overlapping voxels for the two between-group difference maps at a given peak level (i.e. with different levels of thresholding applied). To accomplish this we first rank-ordered all the voxels across the between-group difference PFC rGBC map before and after GSR. Next, we selected voxels that fell only in a given range but that were either at the top or bottom of the range (here the top voxels reflect regions that are higher in SCZ than CON, whereas bottom ranked voxels would reflect regions where CON are higher than SCZ). We specifically used a sliding window where we selected the top/bottom 1%, 5%, 10% and 20% of the voxels. Put simply, if the maps show a large degree of similarity then the spatial overlap for each of the selected ranges should be significantly higher than what would be expected by chance (and should not be dominated by only a single range). After computing the percentage of voxels that overlap across the two between-group difference maps, we tested for significance of overlap using the binomial proportion test (as done in prior work (1)) (see **Fig. S8C**).

Thalamo-Cortical Connectivity Analyses as a Function of Global Signal Removal (GSR). To quantify whether GSR altered thalamo-cortical relationships in SCZ versus BD, we computed two separate but complementary analyses:

i) We quantified the relationship between CGm (i.e. the entire cortical gray matter for each subject) and the thalamus. This analysis was explicitly designed to be non-specific (i.e. considering the entire CGm BOLD signal) to provide proof-of-principle examination of the ways in which large-scale thalamo-cortical relationships can be altered by GSR differentially across different neuropsychiatric conditions. We computed thalamo-cortical connectivity using the Pearson's correlation coefficient (r), which was then de-composed into variance and covariance components. The aim was to examine whether shared versus non-shared sources of variance in cortical/thalamic signals are differentially altered as a function of GSR. Put simply, GSR may uniquely impact the variability of different regional BOLD signals in SCZ. In such a case the correlation coefficient may not be a precise measure given that it combines both shared and non-shared variability(29).

ii) Several studies(30, 31), including our group(1), identified profound alterations in thalamo-cortical connectivity in SCZ. These neuroimaging investigations observed a pattern whereby SCZ was associated with thalamic over-connectivity with sensory-motor regions, but thalamic under-connectivity with prefrontal-cerebellar regions(1, 31). Here we rigorously tested whether GSR would affect between-group differences within areas identified in our prior study. Although we examined the effects of GSR broadly in the analyses described above (juxtaposing the total CGm results before and after GSR), here we explicitly quantified its effects within the identified subset of areas known to show thalamic-cortical dysconnectivity in SCZ. Briefly, we computed the whole-brain seed-based thalamus correlation map using the N=90 SCZ discovery sample, a result originally reported and characterized in our prior study(1) (see **Fig. S6-7**). Next, we extracted the average time-series across all voxels in each subject's bilateral anatomically-defined thalamus (again through subject-specific Freesurfer-based segmentation(20, 32)). This thalamic signal was then correlated with each gray matter voxel and the computed Pearson correlation values were transformed to Fisher Z values (Fz) using a Fisher r-to-Z transform. This yielded a map for each subject, where each voxel's value represents connectivity with the thalamus. Between-group differences, Fz maps were entered into an independent samples t-test, results of which were reported in detail in our prior work(1). We used these identified areas

showing over/under thalamic connectivity in SCZ to test the effects of GSR on the pattern of between-group differences for the discovery SCZ sample (N=90), which served to identify these effects originally(1). We repeated all analyses for the SCZ replication sample (N=71) and BD sample (N=73), both fully statistically independent of the discovery effect (**Fig. S6-7**). The key objective was to examine if GSR would alter the pattern of between-group inferences when focusing on specific thalamo-cortical systems known to exhibit dysconnectivity in SCZ.

Computational Modeling Framework. As noted in the main text, here we employed a validated computational model of resting-state functional connectivity(33), extending a local circuit model(34), to incorporate biologically plausible mechanisms for BOLD signal fluctuations. The network is composed of 66 nodes and is a dynamic mean-field model(34), coupled through a large-scale anatomical connectivity matrix, which was derived from diffusion tractography in humans as reported in Hagmann et al. (35). We extracted the anatomical connectivity matrix from Fig. 1 of Deco & Jirsa(36). Specifically, the digital image provided an RGB value for each connection that we quantitatively matched to a calibrated color bar, allowing us to extract the quantitative strengths of the connections which we used as in prior modeling studies(36).

Of note, the strengths of global connections allow for an undetermined scale factor, which accounts for differences in the range of G and w values between this study and that of Deco et al.(33). BOLD signals were simulated via the Balloon-Windkessel model, as done previously(Deco et al., 2013). We parametrically varied strengths of local self-coupling (w) within nodes and global coupling (G) between nodes. Default values were set to w=0.531, G=1.25, and sigma=0.0004. All reported effects did not depend qualitatively on the specific default values, as shown in **Fig. 5F**. All other parameter values were set to those of Deco and colleagues(Deco et al., 2013). To realistically capture the percentage of BOLD signal variance represented by the GS variance, we introduced a common input to all nodes, mimicking low baseline level correlations seen empirically in gray matter. Specifically, a shared noise term of amplitude 0.0005 was added to the signal of all nodes.

In the original model implementation, Deco and colleagues (2013) utilized different default w/G/sigma values and did not incorporate a common shared signal component. When using the original approach, we found the results to be somewhat unrealistic given the known large contribution of GS in empirical data (see **Fig. 1**). For instance, in our empirical data, we found that ~90% of the variance of the healthy control BOLD signal was accounted for by the GS component, which constitutes the upper bound of GS, reflecting both neuronal and physiological signal present *in vivo*. Therefore, we implemented a common noise input into our model architecture. This way we were able to simulate a neuronal GS that accounts for 77% of the total simulated BOLD signal variance at default values given above (see Figure S15 for a parameter sweep showing percentage of GS at different values of w/G/sigma). We opted for this value to account for the fact that empirical GS certainly contained a physiological non-neuronal component. In turn, these adjustments produced a very robust effect of GSR *in silico* (see **Figure 5**, dotted lines), as seen empirically (see **Figure 1**, dotted lines). Put simply, by including a 'common' noise signal (amplitude=0.0005) we observed a much more realistic effect of GSR in the model.

Implementing GSR in the Model. In the model architecture, we implemented a GSR component in the following way:

i) A GS was computed across as the spatial average over all n=66 nodes:

$$GS(t) = \frac{\sum_i^n BOLD_i(t)}{n}$$

where t is time and $BOLD_i(t)$ is the signal for node i as a function of time.

ii) To mimic the effects of GSR, as computed empirically, we conducted GSR for each node in the model (as would be done per voxel in vivo):

$$BOLD_i^{GSR}(t) = BOLD_i(t) - GS(t)$$

where again t is time, $BOLD_i(t)$ is the signal for node i as a function of time, and $BOLD_i^{GSR}(t)$ is the signal after GSR for that node.

SI RESULTS

Influence of Confounding Variables on Cortical Power & Variance in SCZ. The increase in CGm power and variance is robust across our two independent SCZ samples, but it could still be related to statistical or physiological confounds. We examined several possible confounds suggested by previous literature: i) smoking status, which could affect BOLD fluctuations via oxygen availability in the lungs(37); ii) psychotropic medication levels, which affect measured functional connectivity(38); and iii) head motion(16) (**Fig. S3**).

First, we explicitly compared smoker (N=50) vs. non-smoker (N=46) SCZ patients (collapsing across both replication and discovery samples with available smoking status information, N=96 total). CGm power [t(94)=.51, p=.61, NS] and variance [t(94)=.54, p=.59, NS] did not differ as a function of smoking status (**Fig. S3A-B**). Consistent with the findings not being driven by confounds of smoking, CGm power [t(138)=3.23, p<.0016] and variance [t(138)=3.22, p<.0016] were still significantly higher for non-smoking SCZ patients vs. non-smoking HCS prior to GSR (**Fig. S3A-B**).

To examine the effect of medication, we converted all medication levels to chlorpromazine (CPZ) equivalents, a standard measure of overall level of psychotropic medication(11). For the discovery SCZ sample the CPZ equivalents did not significantly correlate with either CGm power (r=-.03, NS) or variance (r=-.02, NS) (**Fig. S3C-D**). The same result was observed in the replication sample for power (r=.1, NS) and variance (r=.1, NS), suggesting medication dosage cannot explain the GS effects observed in SCZ.

Lastly, we sought to rigorously establish that movement did not drive our core effects via two follow-up analyses. First, we examined the CGm power/variance effects across N=100 HCS and N=100 SCZ (evenly selected across the samples) explicitly statistically matched for the number of frames eliminated during movement scrubbing (19% frames for SCZ and 17.22% frames for HCS, p=.22, NS). We also movement-scrubbed these data for the variance analyses (**Fig. S3F**). The CGm power [(t(198)=2.12, p<.036] and variance [(t(198)=2.14, p<.035] remained elevated in SCZ relative to HCS (**Fig. S3E-F**, left panels). We performed the same analysis in N=73 SCZ and N=73 BD patients statistically matched for the number of frames scrubbed (13.22% frames for SCZ and 11.24% frames for BD, p=.13, NS). This cross-diagnostic analysis further confirmed that increased CGm power [t(144)=2.03, p<.045] and variance [t(144)=2.1, p<.04] are specific to SCZ, even when explicitly comparing movement-matched clinical groups (**Fig. S3E-F**, left panels). In combination, these analyses are not supportive of a relationship between CGm power/variance effects in SCZ and smoking, medication or movement.

Between-Group Differences in Variance and Covariance Structure of Thalamo-Cortical Signals Are Reduced and Change Sign As a Function of GSR. We demonstrate that the average CGm signal power is significantly increased in SCZ and that GSR reduces average CGm power/variance differentially across clinical groups (**Fig. 1**). As noted, this discovery has methodological implications for all future connectivity studies. To illustrate this issue, we present data-driven connectivity analyses (**Fig. 4**) suggesting that removal of large GS variance from one of the groups can affect the final pattern of between-group results. Another possible methodological consequence is that such large GS variance could in turn alter system-level covariance relationships differentially before and after GSR in SCZ(27).

Building on recent evidence suggesting that thalamo-cortical connectivity is altered in SCZ(31, 39), we first investigated the consequences of differential GSR effects upon large-scale thalamo-cortical relationships. Specifically, we examined the correlation (a commonly used measure of functional connectivity) between thalamic and CGm signals with and without GSR. However, use of correlations can at times be problematic in between-group clinical comparison because this measure combines the shared and non-shared sources of variability(29). That is, the correlation coefficient normalizes the covariance between signals, by using the variance of each of the signals (**Fig. S6**, bottom panel). We hypothesized that this can result in conflicting between-group inferences if the variance structure of each signal is different between groups and/or

differentially altered across groups as a function of GSR. Therefore, to fully parse effects, we decomposed the correlation into covariance and variance components (**Fig. S6A-D**).

As noted, CGm signal variance was significantly higher in SCZ relative to HCS [$F(1,178)=7.24$, $p<0.01$]. Also, thalamic and CGm variances were significantly altered across groups as a function of GSR (thalamic variance alteration [$F(1,178)=12.02$, $p<0.001$], CGm variance alteration [$F(1,178)=5.246$, $p<0.03$]). We also observed a significant group difference in CGm-thalamic covariance [$F(1,178)=21.52$, $p<0.0001$] (**Fig. S6A**, top), which GSR altered differentially for each group (i.e. *group x preprocessing* interaction [$F(1,178)=12.52$, $p<0.001$]). Next, by computing the correlation coefficient (**Fig. S6D**), we found in the discovery sample that patients' average correlations between CGm and thalamic signal also significantly differed from HCS (i.e. group main effect [$F(1,178)=23.59$, $p=0.0001$]) (**Fig. S6D**, top). However, the GSR effect on between-group differences was no longer significant (i.e. *group x preprocessing* interaction [$F(1,178)=0.32$, $p=0.57$, ns]). That is, the average thalamo-cortical correlation group differences were similar before and after GSR (**Fig. 6D**, top). This was in contrast to the measures of average CGm-thalamic covariance, where SCZ patients' measures became more similar to those of HCS after GSR. All findings were fully replicated in the independent SCZ sample (n=71) (**Fig. S6A-D**, middle row), which underwent identical movement scrubbing (see **Experimental Procedures**). This might initially suggest that correlations can perhaps be effective in dealing with GSR vs. non-GSR differences when conducting between-group comparisons because between-group correlation differences do not seem to be attenuated by GSR. However, correlations could be potentially problematic in some cases for the same reason that GSR is problematic: both obscure a variance component from between-group comparisons. The observation that GSR has no effect on reducing between-group correlation differences likely arises from correlations having already 'normalized' some of the variance that would otherwise have been removed by GSR.

A parallel examination of the BD sample showed significant between-group differences in thalamic variance (**Fig. S6C**, bottom row) [$F(1,127)=8.99$, $p<0.01$] and CGm-thalamic covariance (**Fig. S6A**, bottom row) [$F(1,127)=7.37$, $p<0.01$], but no significant between-group differences in CGm variance (as reported in **Fig. 1**). In contrast to SCZ, GSR did not alter between-group differences in the BD sample for any of the measures, again demonstrating diagnostic specificity of the SCZ GSR effect (all p-values >0.08). Interestingly, the thalamo-cortical covariance and correlation relationships for the BD patients show striking similarities to those found for SCZ patients when contrasted with HCS. Yet, BD exhibited lower CGm and thalamic variance than SCZ prior to GSR. In other words, the correlation structure may be insufficient to distinguish SCZ from BD; more specificity was gained from additionally examining variance. Collectively, these observations further suggest that correlations may sometimes not be ideal when the variance structure drastically differs across clinical groups. Alternatively, depending on the goal of the study, correlations could be used because they provide results untarnished by differences in variability. Considering these pros and cons will likely depend on the type of research question and selecting the best statistical metric for a given inference.

Notably, GSR did not significantly alter the between-group differences of CGm-thalamic correlation for any of the patient samples (**Fig. S6D** top, middle, bottom rows). Yet, GSR did significantly alter the magnitude of between-group differences in CGm-thalamic covariance for both SCZ samples but not for the BD sample. In addition, across all analyses, GSR resulted in a 'shift' of positive thalamo-cortical covariance (or correlations) into the negative range for all three groups, altering interpretations for all findings, as predicted by prior simulation studies(40). That is, in each case SCZ and BD showed a more positive value relative to HCS, but after GSR all values changed sign and the gap between SCZ and HCS relationships reduced – a notable difference from the inferred connectivity strength before GSR. Collectively, these effects show how GSR can in some cases impact the variance and covariance structure of large-scale thalamo-cortical systems for different between-group comparisons, leading to possibly conflicting inferences depending on the preprocessing method.

Covariance/Correlation Structure of A Priori Defined Thalamo-Cortical Networks Do Not Change Sign After GSR. As noted above, based on the discovery of larger GS power/variance, an important objective was to explore how GSR could differentially impact inferences across different clinical analyses. We first focused on the relationship between thalamus and CGm, which is known to be disrupted in SCZ. As discussed above, these covariance analyses were significantly affected by GSR. However, GSR could potentially produce

different effects when focusing on thalamo-cortical relationships within *a priori* selected regions(31) of cortex, rather to the entire cortex. To test this more specific case we examined previously identified regions that show thalamo-cortical dysconnectivity in SCZ, independently replicated by several groups(1, 30, 31). Specifically, prior studies identified two types of thalamo-cortical alterations in SCZ: i) regions showing reductions in thalamic coupling (i.e. under-connectivity), and ii) regions that show increases in thalamic coupling (i.e. over-connectivity)(1). Here we tested, using these *a priori* defined areas (see **SI Experimental Procedures** above), whether GSR altered between-group inferences. This analysis accomplishes two important additional objectives: i) based on present effects it is vital to examine their methodological implications; namely, whether GSR alters established findings in the literature; ii) this analysis provides another replication of prior effects(1), via an independent SCZ sample (N=71). In turn, such independent verification of thalamo-cortical effects provides, irrespective of GSR, additional confidence for future studies focusing on this finding as a viable SCZ biomarker.

Notably, these analyses are distinct from results above (which averaged the entire CGm signal). That is, here we examine whether between-group differences in subsets of thalamo-cortical connections remain unaltered by GSR. As predicted, we observed significant GSR effects(40). However, correlation and covariance measures for SCZ patients retained the same sign after GSR (**Fig. S7**) instead of inverting as observed for the total CGm analysis (**Fig. S6**). HCS patterns generally retained the same sign after GSR (**Fig. S6**). Thus, the overall polarity of thalamo-cortical findings (irrespective of GSR) remained unchanged across samples, showing that certain types of between-group analyses remain more robust to GSR. The discrepancy between these system-specific effects and analyses above that focused on the entire cortex could occur if GSR homogeneously decreases variance across cortex in SCZ (described in **Fig. S3**). This would effectively shift the whole-brain pattern of covariance relationships in a way that attenuates differences between patients and HCS (since GSR drastically reduces a major difference between SCZ and HCS: the CGm variance). Conversely, when *a priori* regions are defined for functional connectivity (as done here using prior discoveries(1)), then the between-group effect may remain robust and retain the same sign irrespective of GSR. This consideration is potentially important for studies focusing on *a priori* selected regions versus those using data-driven analyses (see **Fig. 4 and S8** for data-driven effects focused on PFC).

Lastly, here we also found that using the correlation coefficient did not always fully capture the alteration in thalamo-cortical relationships before and after GSR (as was observed for covariance, see **Fig. S6**). This difference of sensitivity to GSR is caused by the large variance component in the denominator of the correlation equation for SCZ prior to GSR – the resulting correlation “divides out” the variance, and is therefore less sensitive to further decrements in variance resulting from GSR, whereas covariance is more vulnerable to variance-reducing processes such as GSR. Collectivity, these thalamo-cortical effects illustrate a complex impact of GSR on system-level inferences in SCZ (see **Discussion** for a set of recommendations for future clinical connectivity studies).

Computational Modeling – Calculating Mean GS Values. In addition to variance calculations, we also examined mean levels of node-wise signal amplitude and GS amplitude in the model. We found that the mean of simulated GS also increased as a function of G and w (**Fig. S10**). This pattern was expected because the mean value of each simulated local signal also increased as a function of G and w, and GS is the average of local signals. This increase in mean signal values reflects an increase in local recurrent excitation, which will in turn drive up the signal at each local node. Additionally, increased global coupling across nodes will result in an elevated signal propagating throughout all the nodes, thereby increasing mean local node signal levels.

Computational Modeling – Examining Effects of GSR. Once implemented, GSR resulted in attenuated local variance in the model, a pattern that was quite similar to the empirical effects (see **Fig. 1**). However, the GS variance was virtually completely effaced given that implementing GSR in the model subtracts the entire mean of the model-derived signal across all time points. What remains is mathematically zero (see **Fig. S11**).

SI DISCUSSION

Implications for the controversy regarding the role of global signal regression in resting state connectivity analyses. The principal findings of the current study do not offer an unequivocal suggestion to perform GSR or not. However, several attempted rigorous analyses (see **Fig. 4** in the main text, as well as **Fig. S6, S7 & S8**) demonstrate that this is complex issue, depending on the particular analysis. Therefore, the clearest possible recommendation that one can make to the neuroimaging community based on these results is to carefully examine and analyze the GS across all clinical connectivity analyses. As noted in the main text, these effects indicate that GS needs to be characterized explicitly in clinical groups to determine its potential contributions in focused connectivity analyses (see **Fig. S6-7**). Only once such analyses have been completed in a given dataset, researchers can reach a more informed decision if GSR is advisable. Furthermore, as shown here, future studies should test for the direction of clinical inferences before and after GSR to allow a nuanced interpretation regarding the observed connectivity alterations. Such a step-wise approach is critical to improve connectivity-derived neuroimaging markers in neuropsychiatric research, circumventing the debate whether to use GSR or not; instead employing rigorous data inspection to support appropriate study-specific analytic decisions.

SI TABLES & FIGURES

Table S1 | Clinical and Demographic Characteristics

Characteristic	HCS (N=90)		SCZ (N=90)		Significance	
	M	S.D.	M	S.D.	T Value / Chi-Square	P Value (two-tailed)
Age (in years)	30.71	11.99	32.93	11.25	1.28	0.2
Gender (% male)	66.00		73.00		1.13	0.26
Father's Education (in years)	14.37	3.21	13.67	3.47	1.42	0.16
Mother's Education (in years)	13.99	2.81	13.50	2.92	1.15	0.25
Participant's education (in years)	15.24	2.22	13.18	2.21	6.26*	<.001
Handedness (% right)	84.21		80.00		0.85	0.4
Signal-to-noise (SNR)	215.37	45.25	206.81	62.05	1.06	0.3
% Frames Flagged	10.13	7.89	17.63	17.00	3.79*	<.001
IQ Estimate	106.77	8.92	97.78	15.71	4.55*	<.001
Medication (CPZ equivalents)	-	-	229.00	195.81	-	-
PANSS Positive Symptoms	-	-	15.80	4.73	-	-
PANSS Negative Symptoms	-	-	14.34	5.53	-	-
PANSS General Psychopathology	-	-	30.48	7.18	-	-
PANSS Total Psychopathology	-	-	60.51	14.25	-	-

Table S1. Discovery Schizophrenia Sample Demographics. HCS, Healthy Comparison Subjects; SCZ; Patients diagnosed with Schizophrenia; PANSS, Positive and Negative Syndrome Scale; M, Mean; SD, Standard Deviation; IQ, intelligence quotient; age, education levels, parental education, are expressed in years. CPZ, Chlorpromazine equivalents were calculated according to latest validated approaches(11). * denotes a significant T statistic for the between-group t-test.

Table S2 | Clinical and Demographic Characteristics - Schizophrenia Replication Sample

Characteristic	HCS (N=74)		SCZ (N=71)		Significance	
	M	S.D.	M	S.D.	T Value / Chi-Square	P Value (two-tailed)
Age (in years)	35.82	11.58	38.14	13.99	1.09	0.28
Gender (% male)	69.00		80.00		1.57	0.12
Parental Education (in years)	4.62	1.83	4.15	2.05	1.45	0.15
Participant's education (in years)	4.64	1.31	3.93	1.43	3.12*	<.01
Handedness (% right)	93.00		80.00		1.91	0.06
Signal-to-noise (SNR)	213.38	79.79	199.47	93.62	0.97	0.33
% Frames Flagged	18.81	20.83	31.10	28.63	2.98*	<.01
Medication (CPZ equivalents)	-	-	372.53	304.82	-	-
PANSS Positive Symptoms	-	-	14.85	4.76	-	-
PANSS Negative Symptoms	-	-	14.52	4.86	-	-
PANSS General Psychopathology	-	-	29.15	8.38	-	-
PANSS Total Psychopathology	-	-	58.52	13.76	-	-

Table S2. Replication Schizophrenia Sample Demographics. HCS, Healthy Comparison Subjects; SCZ; Patients diagnosed with Schizophrenia; PANSS, Positive and Negative Syndrome Scale; M, Mean; SD, Standard Deviation. Education level for the replication sample was determined based on the following scale: Grade 6 or less=1; Grade 7–11=2; high school graduate=3; attended college=4; graduated 2 years college=5; graduated 4 years college=6; attended graduate or professional school=7; Completed graduate or professional school=8. CPZ, Chlorpromazine equivalents were calculated according to latest validated approaches(11). * denotes a significant T statistic for the between-group t-test.

Table S3 | Clinical and Demographic Characteristics - Bipolar Sample

Characteristic	HCS (N=56)		BD (N=73)		Significance	
	M	S.D.	M	S.D.	T Value / Chi-Square	P Value (two-tailed)
Age (in years)	31.25	10.35	32.00	11.27	0.39	0.70
Gender (% male)	38.71		27.03		1.84	0.07
Paternal education (in years)	12.98	3.87	14.83	3.59	2.79*	0.01
Maternal education (in years)	13.63	2.58	13.99	2.60	0.78	0.44
Participant's education (in years)	15.11	2.10	14.22	1.90	2.51*	0.01
Signal-to-noise (SNR)	215.45	58.98	216.80	52.03	0.89	0.14
% Frames Flagged	9.74	10.44	11.24	9.78	0.41	0.84

Table S3. Bipolar Disorder Sample Demographics. M, Mean; SD, Standard Deviation; age, education levels, and parental education are expressed in years. *denotes a significant T statistic for the between-group t-test. For complete clinical details and clinical measures used for the bipolar sample please see prior studies(7).

FreeSurfer Code		Landmark
		Whole-brain Labels
Left	Right	
3	42	Cerebral Cortex - entire hemisphere
8	47	Cerebellar Cortex - entire hemisphere
10	49	Thalamus
11	50	Caudate
12	51	Putamen
13	52	Pallidum
17	53	Hippocampus
18	54	Amygdala
26	58	Accumbens
28	60	Ventral DC
16	16	Brainstem
		Prefrontal Cortex Labels
Left	Right	
1002	2002	Caudal Anterior Cingulate Cortex
1003	2003	Caudal Middle Frontal Cortex
1012	2012	Lateral Orbitofrontal Cortex
1014	2014	Medial Orbitofrontal Cortex
1018	2018	Inferior Frontal Cortex - Pars Opercularis
1019	2019	Inferior Frontal Cortex - Pars Orbitalis
1020	2020	Inferior Frontal Cortex - Pars Triangularis
1026	2026	Rostral Anterior Cingulate Cortex
1027	2027	Rostral Middle Frontal Cortex
1028	2028	Superior Frontal Cortex
1032	2032	Frontal Pole

Table S4. List of FreeSurfer Codes Used for Gray Matter, Thalamus and PFC-restricted rGBC Analyses.

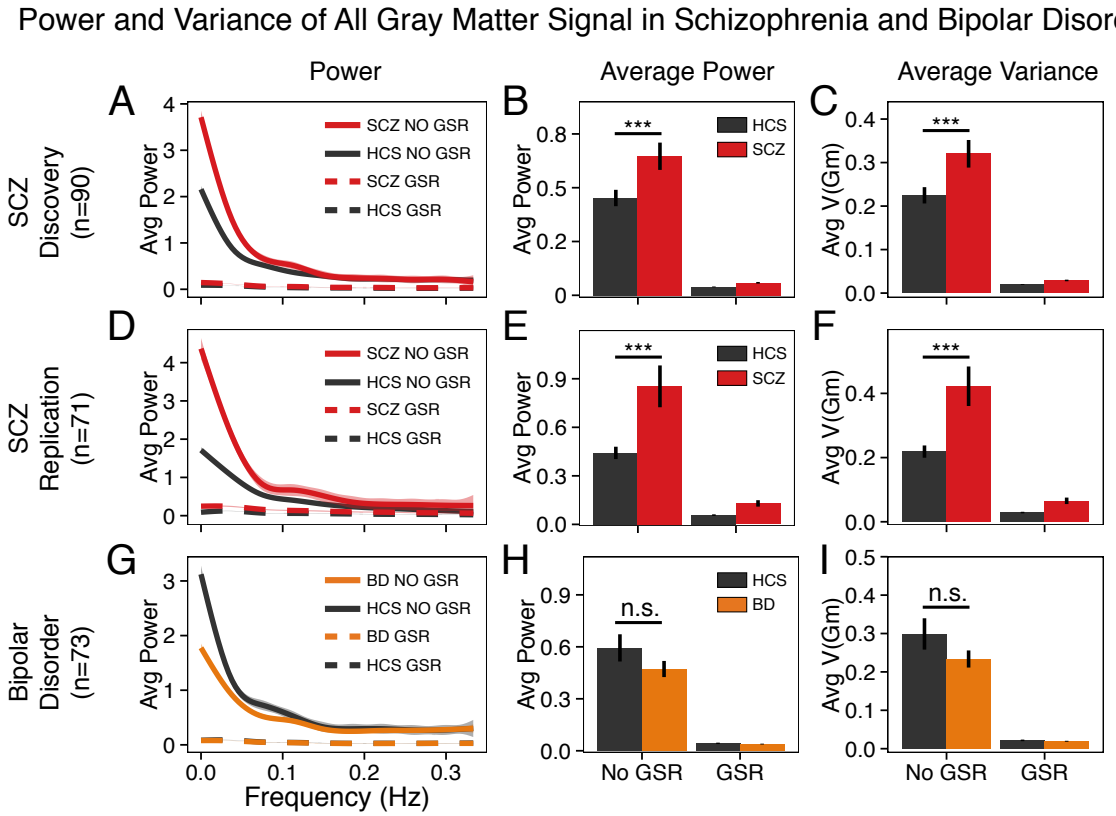

Fig. S1. Power and Variance of All Gray Matter Signal in Schizophrenia and Bipolar Disorder. (A) Left: Power of the gray matter signal (Gm) in 90 schizophrenia (SCZ) patients (red) relative to 90 healthy control subjects (HCS) (see **Table S1** for demographic details) extending cortex-specific effects (**Fig. 1**). (**B**) The bar graph shows the mean power across the entire frequency range for both groups before and after GSR, indicating an increase in SCZ [$F(1,178)= 7.64$, $p<0.01$]. As with the cortex analyses, the effect is significantly attenuated after the GSR processing step is used [$F(1,178)= 5.9$, $p<0.02$]. (**C**) Variance of the Gm replicated the power effects showing increases in SCZ [$F(1,178)= 7.42$, $p<0.01$] and also verified that GSR attenuates the finding [$F(1,178)= 5.75$, $p<0.02$]. (**D-F**) As with the cortical analyses, here we fully replicated the pattern of results using an independent sample collected at a different site and diagnosed independently (see **Table S2** for demographic details), again confirming the increase in Gm power [$F(1,178)= 10.8$, $p<0.0015$] and variance [$F(1,143)=11.44$, $p<0.001$] effects in SCZ. The replication sample also verified the attenuating impact of GSR preprocessing on power [$F(1,143)= 8.34$, $p<0.01$] and variance [$F(1,143)= 8.99$, $p<0.01$] metrics in SCZ vs. HCS. (**G-I**) The results for bipolar disorder (BD) patients (N=73) relative to demographically matched HCS (see **Table S3** for BD demographic details) did not reveal the effect of GSR observed in the SCZ samples [$F(1,127)=3.05$, $p=0.08$, n.s.] and there was no evidence for increase in Gm power or variance for the BD patient sample. These results across all gray matter voxels follow the effects reported for the cortical gray matter analyses (see **Fig. 1**). Error bars mark +/- 1 standard error of the mean.

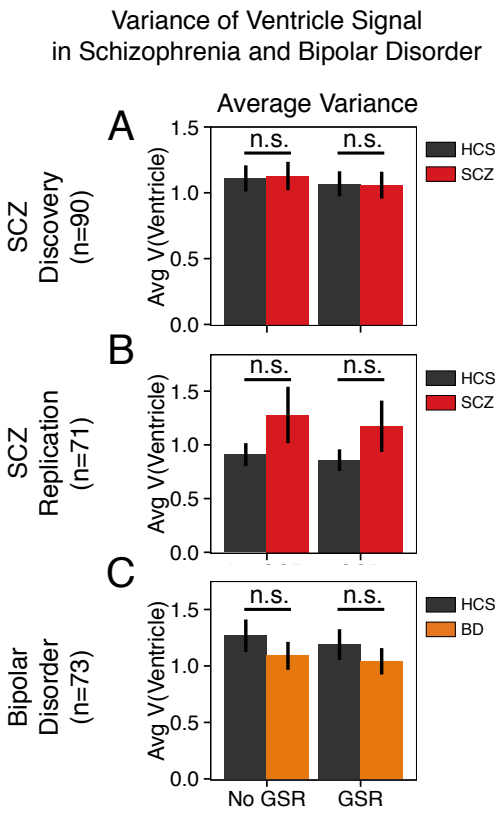

Fig. S2. Increased Power/Variance is Not Present in Ventricles. To examine whether the increase in variance was present in areas without any plausible neurobiological origin, we examined ventricle signal across the discovery and replication schizophrenia (SCZ) samples and the bipolar disorder (BD) sample. We also investigated if GSR has an effect on this non-neuronal signal. Unlike the results presented for cortex and total gray matter, here there was no main effect of diagnosis: (**A-C**) **SCZ-Discovery:** F(1,178)= 0.01, p=0.99; **SCZ-Replication:** F(1,143)= 1.65, p=0.2; **BD:** F(1,127)= 0.74, p=0.4] or diagnosis x preprocessing interaction across any of the samples [**SCZ-Discovery:** F(1,178)= 2.1, p=0.15; **SCZ-Replication:** F(1,143)= 3.24, p=0.08; **BD:** F(1,127)= 2.04, p=0.16]. All ventricle signal for these analyses was defined anatomically for each subject via automated FreeSurfer segmentation(20) and inspected for quality assurance by a trained rater (AA).

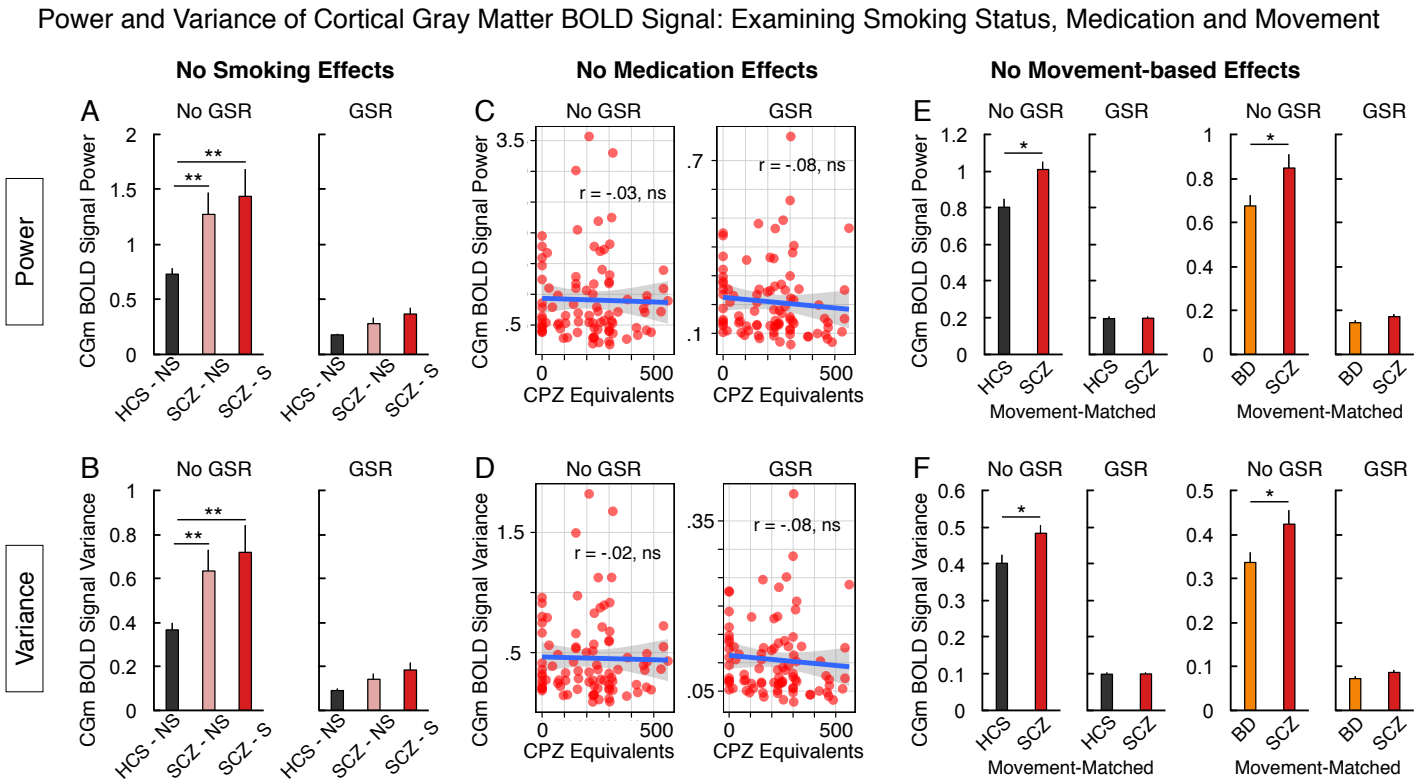

Fig. S3. Examining Confounding Variables: Smoking, Medication and Movement. (**A-B**) Power and variance of the cortical gray matter (CGm) signal is shown for healthy comparison subjects (HCS) who were not smokers (N=90), patients with schizophrenia who were not smokers (SCZ-NS, N=50) and patients with schizophrenia who were identified as smokers (SCZ-S, N=46) across both discovery and replication datasets. The power and the variance of the CGm were significantly increased for both SCZ-NS and SCZ-S relative to HSC prior to GSR (p<.01 for both measures), whereas there was no difference between SCZ-NS and SCZ-S [t(94)=.52, p=.61, NS], suggesting smoking status does not explain the increased CGm variance and power effects in SCZ. Of note, smoking related to lifetime smoking status (i.e. whether a patient was ever a smoker). Therefore, cumulative effects of smoking could not drive effects for non-smokers as they have never smoked. (**C-D**) The magnitude of CGm power and variance before and after GSR did not significantly relate to medication levels across subjects (all p-values >.4), calculated via CPZ equivalent conversion for the discovery sample(11) and did not differ between medicated versus unmedicated patients for the discovery sample (all p-values >.2). The same pattern was present for the replication sample (all p-values >.25). (**E-F**) We identified N=100 HSC and N=100 SCZ evenly across the two samples that were explicitly matched on the amount of frames flagged for movement (19% frames for SCZ and 17.22% frames for HCS, p=.22, NS). The power and the variance of the CGm are significantly increased for SCZ relative to HSC prior to GSR (p<.04 for both measures) even when explicitly matched for number of flagged frames. Next, we identified N=73 SCZ and N=73 BD that were explicitly matched on the amount of frames flagged for movement (13.22% frames for SCZ and 11.24% frames for BD, p=.13, NS). Again, CGm power and variance were significantly increased for SCZ relative to BD prior to GSR (p<.045 for both measures), further establishing the diagnostic specificity of present effects. This analysis in particular guarantees that effects persist when, by definition, they cannot be confounded by movement. That is, head motion cannot confound effects in an analysis where the number of flagged frames is explicitly matched across groups, guaranteeing that the amount of movement is not different across samples. Error bars mark +/- 1 standard error of the mean.

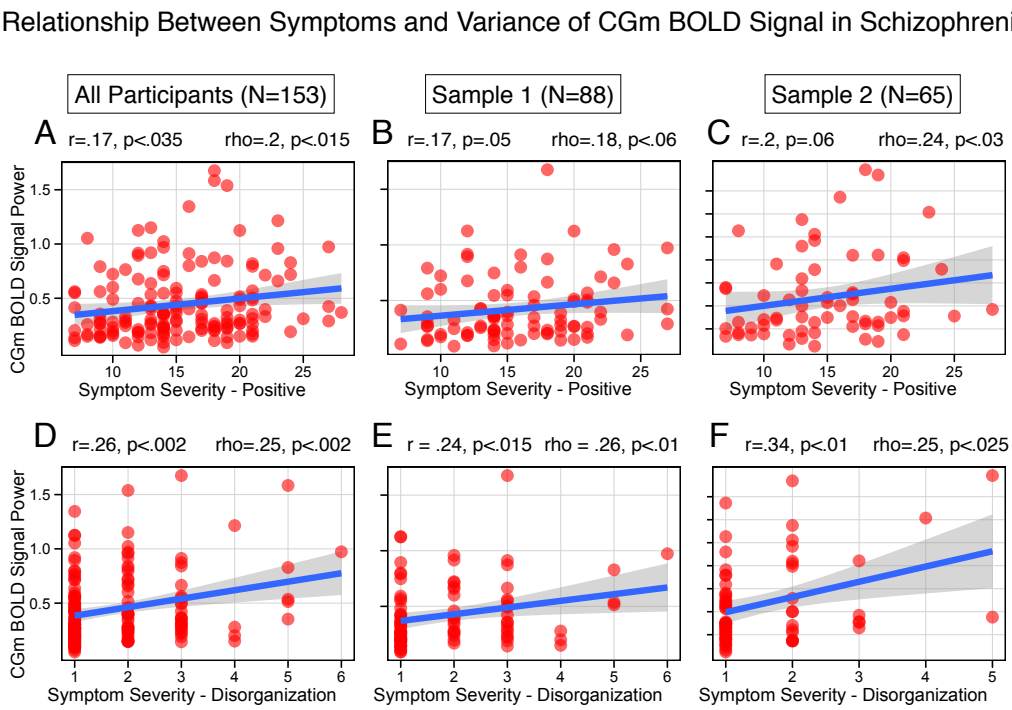

Fig. S4. Relationship Between Schizophrenia Symptoms and Cortical Gray Matter (CGm) Variance. We extracted the average magnitude of the CGm variance for each subject across both samples for which we had available symptom ratings (N=153), as done for CGm power in **Fig. 3**. (**A-F**) All the effects remained unchanged for CGm variance. As with CGm power, all of the relationships with symptoms were again attenuated and no longer significant after GSR for overall positive symptoms as well as disorganization specifically. The shaded area marks the 95% confidence interval.

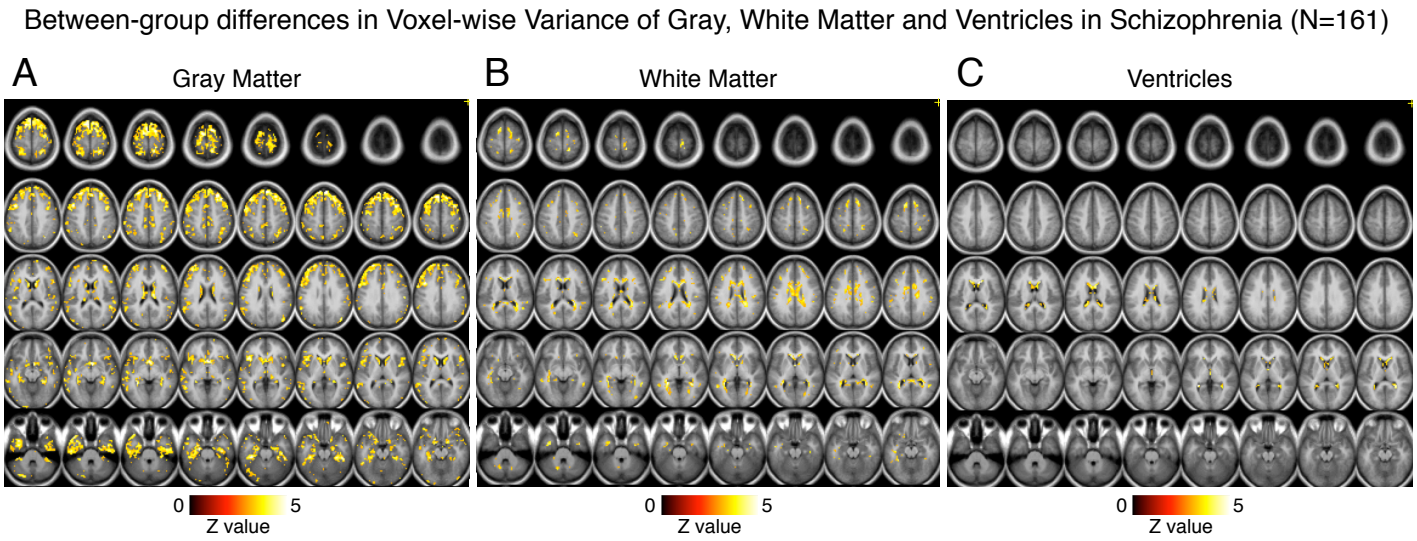

Fig. S5. Increased Voxel-wise Variability Across Gray Matter, White Matter and Ventricles Before GSR. We show voxel-wise map of between-group differences whereby each yellow-orange voxel indicates greater local variability for SCZ relative to matched healthy control subjects (HCS) (at Z>3, p<.0013 uncorrected). This pattern is shown for (**A**) gray matter, (**B**) white matter and (**C**) ventricles. Notably, the increase in voxel-wise variance was clearly evident in gray matter, but was also somewhat apparent across other tissue types. One possibility, despite the careful implementation of movement scrubbing(16),(41) is that these patterns (especially in white matter and ventricles) are in part influenced by sub-TR motion (i.e. micro-movements) that remain challenging to fully rule out in the present study, but may propagate as complex waveforms throughout the BOLD signal(16). More careful de-noising and movement correction strategies will be needed to comprehensively remove any possibility of movement-related artifacts contributing to these voxel-wise effects. In addition, using accelerated BOLD acquisition methods(42) could further attenuate the possible impact of motion.

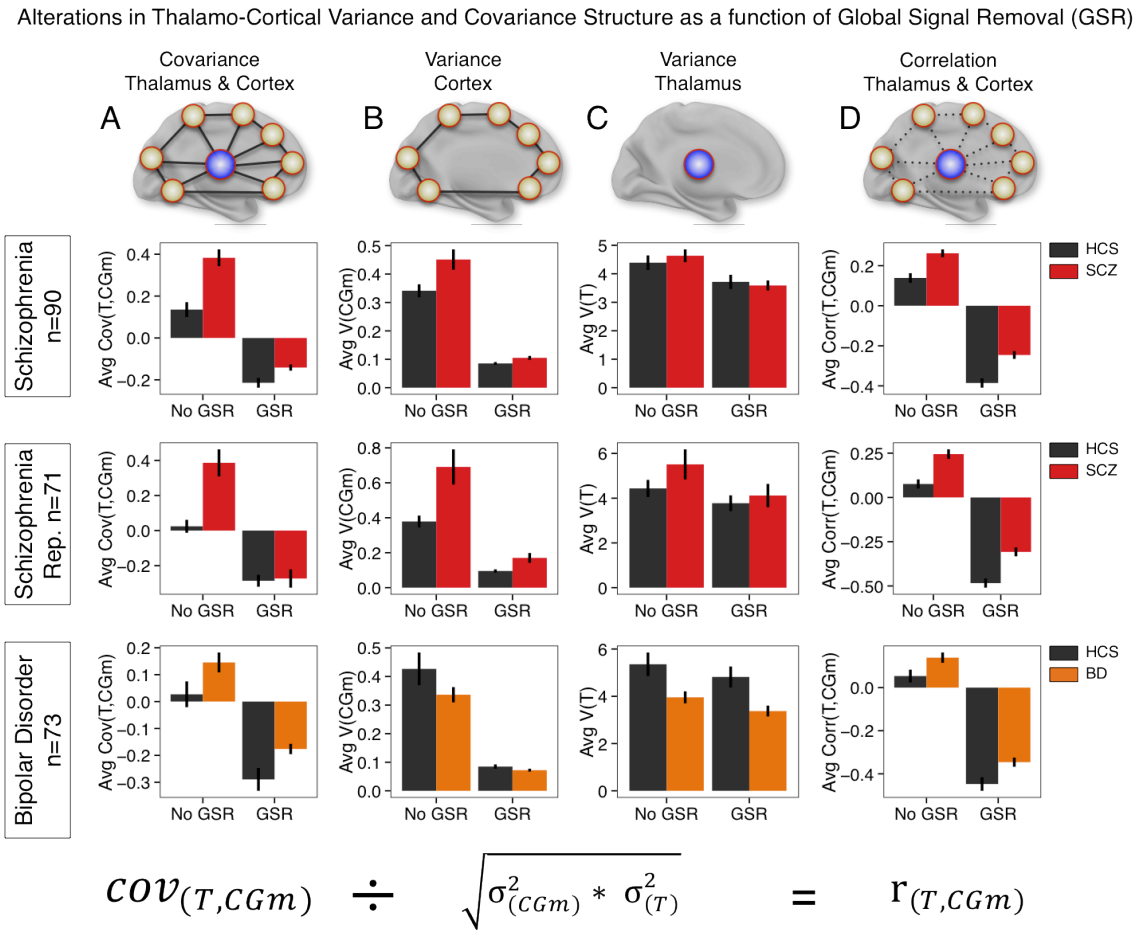

Fig. S6. Thalamo-Cortical Variance & Covariance Alterations as Function of GSR in Schizophrenia and Bipolar Disorder. We found increased GS power/variance in schizophrenia (SCZ) but not bipolar disorder (BD). Given this large variance component in SCZ, we examined if the relationships between thalamo-cortical signals, known to be disrupted in SCZ(1), alter as a function of GSR. We also examined if this large variance component alters the correlation (as opposed to covariance) because correlations would normalize this large variance across groups(see (29) for a detailed discussion). (**A**) Covariance between cortical gray matter (CGm) signal and thalamic signal for the SCZ sample and matched HCS (top), the replication SCZ dataset (middle) as well as BD patients and their matched HCS (bottom). Both results are shown prior to and post GSR. There was a larger between-group difference in the covariance between CGm and thalamic signals seen prior to GSR in SCZ but not BD. (**B**) Variance of the CGm signal is shown. (**C**) Variance of the thalamic signal is shown. (**D**) The correlation between CGm and thalamic signal is shown. For illustrative purpose, figures are presented along with the equation on the bottom to highlight the importance of carefully de-composing the final correlation into variance and co-variance components, given that inference can change. Collectively, these analyses show that removal of GS reverses the sign of the between-group effects, but does not drastically alter the final inference (i.e. SCZ > HCS in thalamo-cortical coupling). However, SCZ is better distinguished from BD in variance comparisons than in correlation or covariance comparisons. Error bars mark +/- 1 standard error of the mean.

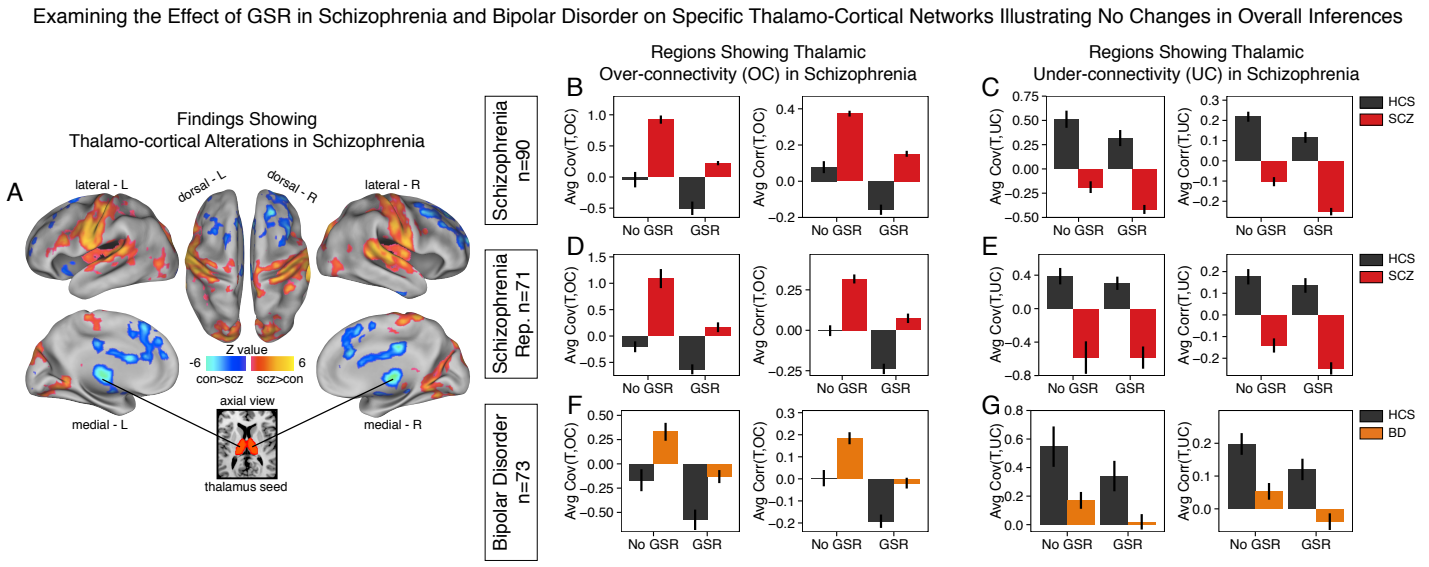

Fig. S7. Examining the Effect of GSR in Schizophrenia and Bipolar Disorder on Specific Thalamo-Cortical Networks. Above we characterized thalamo-cortical signals across the entire cortical mantle (**Fig. S5**); however, we also tested whether GSR alters specific thalamo-cortical network inferences given the large GS power/variance in schizophrenia (SCZ). For this purpose we conducted a focused re-analysis of our previously published effects(1), where we identified both reductions and increases in thalamo-cortical connectivity in SCZ (panel A adapted with permission from (1)). (A) Regions showing reduced (blue, top panel) and increased (red, bottom panel) thalamic connectivity in individuals with SCZ relative to HCS, which we identified in our prior work(1). Because these robust effects were replicated by three independent research groups(1, 31, 43) they provide a basis for investigating the effect of GSR in SCZ, motivated by our primary findings. Here we show results, specifically within the identified areas showing increased (left, panels B,D,F) versus reduced (right, panels C,E,D) thalamic connectivity. Each effect is presented with and without GSR. We conducted this analysis for the original SCZ discovery sample (**B-C**), for the SCZ replication sample (**D-E**) and the bipolar disorder (BD) sample (**F-G**). Collectively, the results illustrate that, irrespective of GSR, the pattern of between-group inferences is not altered and overall findings do not change sign after GSR. Error bars mark +/- 1 standard error of the mean.

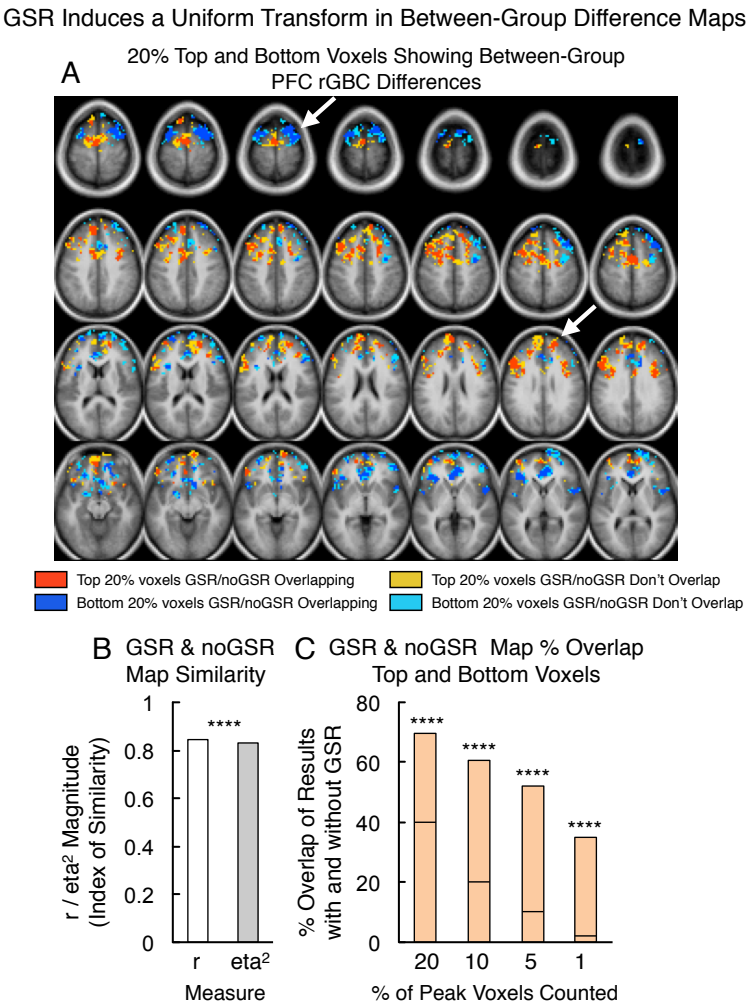

Fig. S8. GSR Induces Uniform Transform of PFC rGBC Between-group Effects. The data-driven rGBC results (**Fig. 5**) suggest that GSR may induce a non-uniform transformation of between-group differences by 'shifting' some connectivity values differentially for schizophrenia (SCZ) vs. healthy comparison subjects (HCS). Here we tested whether results are consistent with this possibility, or whether they are more indicative of a uniform transformation (see **Fig. 5E-F** for a schematic illustration). (**A**) The between-group PFC rGBC difference map is shown for only 20% of the top ranked vs. 20% of the bottom ranked voxels. Dark blue and orange voxels highlight the results with GSR performed, whereas the light blue and yellow voxels highlight the results without GSR performed. As indicated by this qualitative map, the overlap was generally quite high for both positive and negative connection ranges (white arrows highlight two example areas with high overlap). This overlap map is not supportive of a non-uniform transformation. (**B**) To verify this observation statistically we quantified the statistical similarity for the PFC rGBC between-group difference maps before and after GSR using r and eta^2 metrics (we used both measures to provide convergent validity via complementary measures, see **SI Experimental Procedures** for details)(28). Both metrics showed highly significant levels of similarity (both >.8). This robust level of similarity pre and post GSR is inconsistent with the GSR introducing a strictly non-uniform transform (as the similarity would be around chance or approaching 0). (**C**) We also examined the resulting voxel overlap without a threshold applied before and after GSR. We did so across four different peak levels (1%, 5%, 10% and 20% top and bottom ranked voxels showing between-group differences). Again, if GSR fundamentally altered the structure of between-group differences then we would expect the overlap to be quite low across these peak levels. In contrast, we observed that, irrespective of the peak level used, the maps before and after GSR showed a highly significant level of overlap (binomial test for proportions, all p<.0001, chance marked with horizontal lines). Collectivity, these analyses are inconsistent with the hypothesis that GSR induces a non-uniform transform.

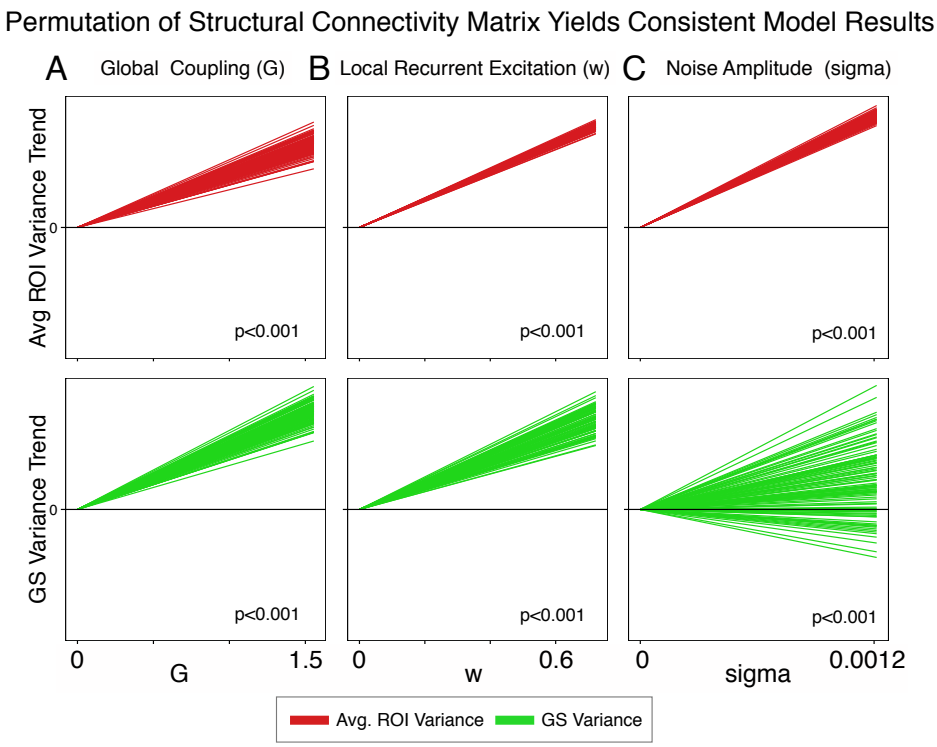

Fig. S9. Permutations of the Structural Connectivity Matrix Does Not Alter the Net Effect Across Parameters. We tested whether the specific configuration of the structural connectivity matrix(35) affected the net pattern of the modeling results. That is, it is important to establish that the predominant effect is an increase in global and local variance as function of G, w and sigma (noise parameter in the model), as opposed to a decrease. To this end, we re-permuted the connectivity matrix randomly (n=100) and re-ran all modeling simulation. (**A-C**) We computed the average node variance (top panels in red) and the global signal (GS) variance across all nodes (bottom panels in green) as a function of G, w and sigma parameters. Next, we computed the slope of the effect for each of the permutations (either positive slope or negative slope). By chance alone we would expect 50% of the modeling simulations to show a positive versus negative slope. However, for each parameter, irrespective of the structural connectivity matrix configuration, we observed a positive slope (binomial test for proportions, p<.001 for each test). This suggests that the modeling results are generic properties of the model and are not necessarily highly sensitive to the particular configuration of structural connectivity patterns.

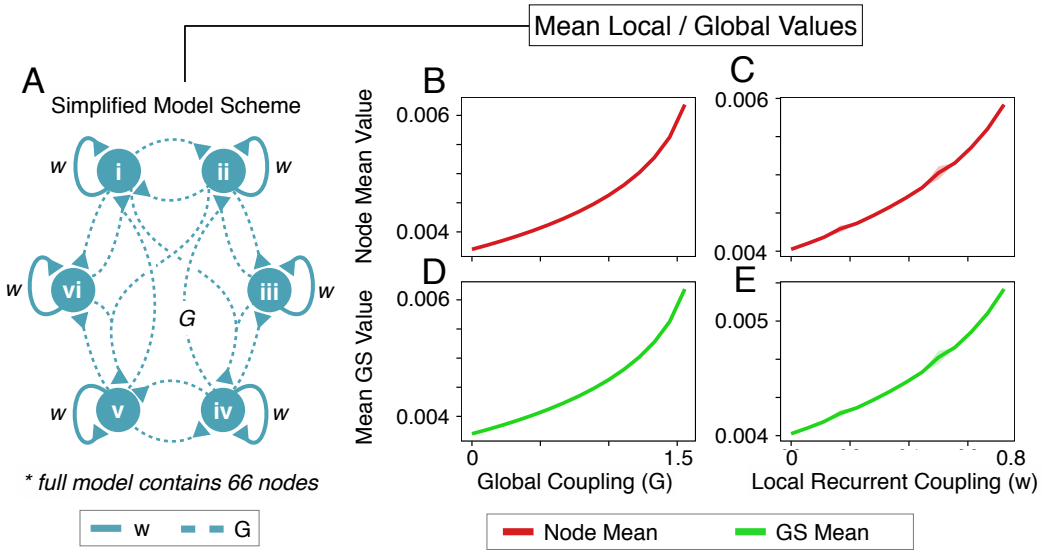

Fig. S10. Calculating Mean GS Values as a Function of w and G. (**A**) The model scheme is shown as in the main text. Here we extracted the mean value calculated for each node and the mean value of the model-derived global signal (GS) (**B-C**) Simulations indicate increased mean of local BOLD signals originating from each node, in response to increased w or G. (**D-E**) The GS, computed as the spatial average across all nodes, also showed increased mean value by elevating w or G. Error bars represent the standard deviation at each value of w or G computed across 4 realizations of the background noise input, illustrating model stability. Here the error bars are not visible due to the small scale.

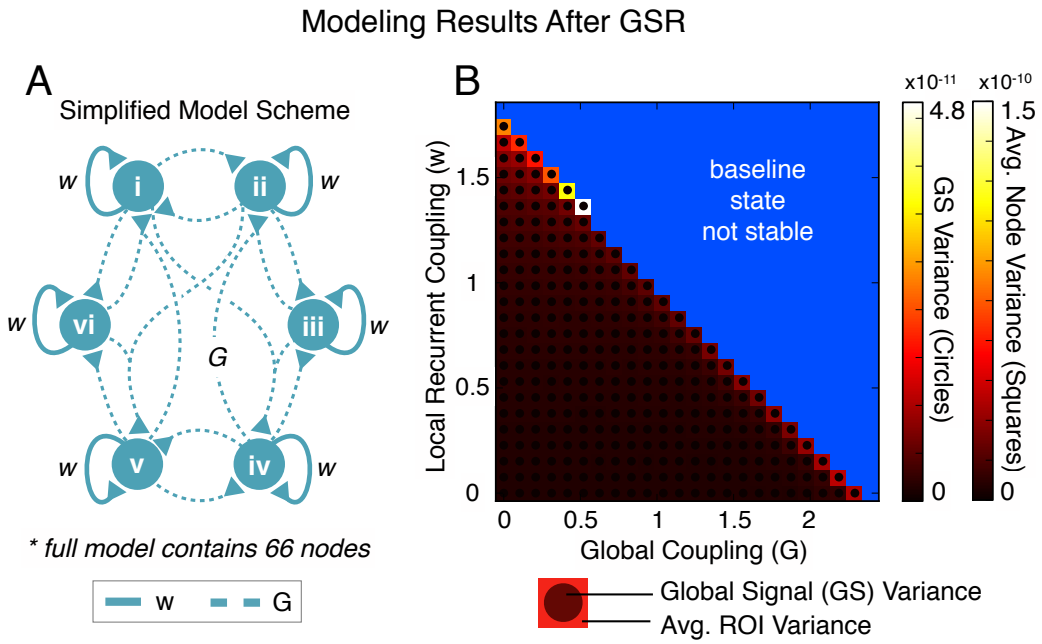

Fig. S11. Quantifying Effects of Simulated GSR. In the main text we presented the effects of *in silico* GSR, juxtaposed next to the original model simulations (see **Fig. 5**). Here we present the full pattern of results following model-based GSR via the parameter-scape visualization. (**A**) The model scheme is shown as in the main text. (**B**) Two-dimensional parameter space following GSR. This simulation captures the decreased variance following GSR at the local node-wise level (squares, far right color bar), Implementing GSR reduces the GS to zero; this is not surprising as GSR effectively subtracts the GS component (circles in each square, the adjacent color bar).

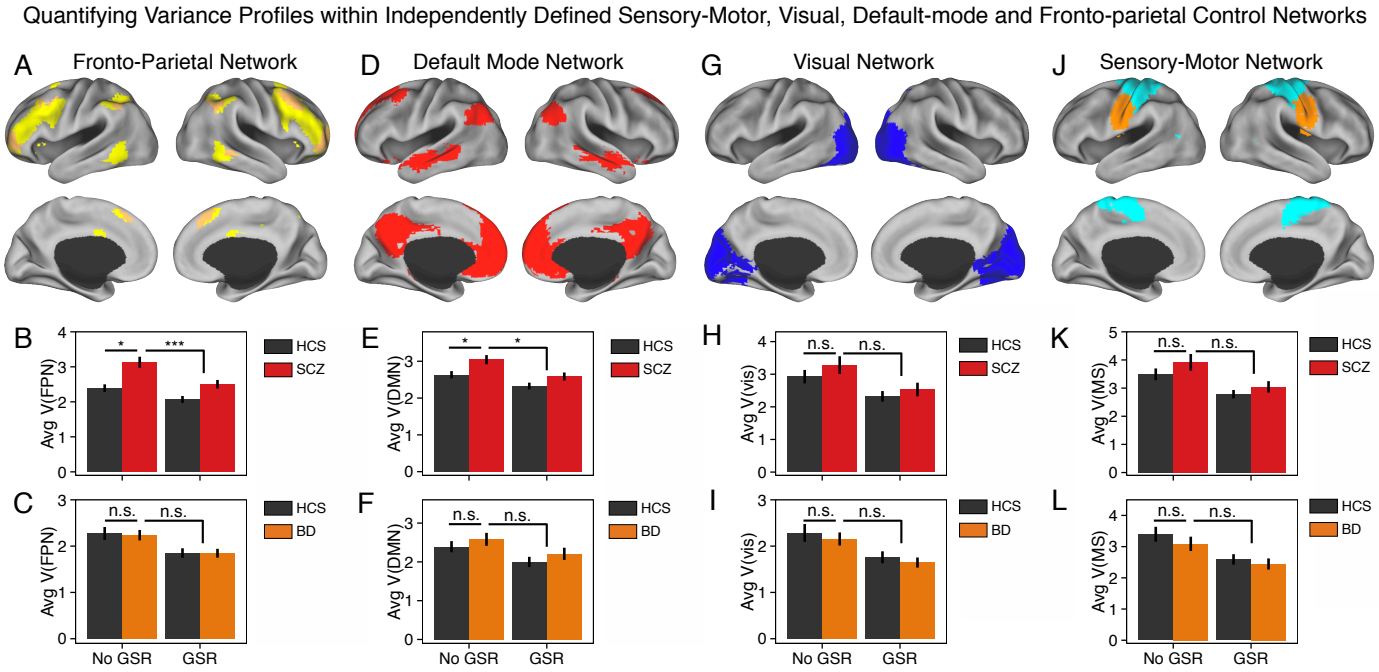

Figure S12. Characterizing Specificity of the Increased Cortical Variance Effects in SCZ – Focus on Large Scale Networks. We discovered significantly elevated GS variability in SCZ (**Fig. 1**). However, it remains possible that this GS variability is preferentially localized to certain networks, especially given existing theoretical models in SCZ suggesting more profound neural disruptions in higher-order associative networks(44). To test this hypothesis, we formally quantified the pattern of elevated variability within specific *a priori* defined large-scale networks, which were characterized independently via resting-state in healthy adults (obtained with permission from Power and colleagues(45)). (**A**) Effects within the fronto-parietal control network. (**B**) SCZ results revealed a significant *Group x Processing* interaction [$F(1,323)=26.74$, $p<4.09x10^{7}$], indicating elevated variance in SCZ prior to GSR [$t(323)=3.88, p<0.00015$], but significantly reduced variance after GSR in SCZ. **(C)** No such effect was evident in BD [$F(1,127)=0.21$, $p=0.65$]. (**D**) Effects within the default-mode network. (**E**) SCZ results revealed a significant *Group x Processing* interaction [$F(1,323)=9.17$, $p<0.003$], indicating elevated variance in SCZ prior to GSR [$t(323)=2.55, p<.015$], but significantly reduced variance after GSR in SCZ. **(C)** No such effect was evident in BD [$F(1,127)=0.07$, $p=0.8$]. (**G-I**) Although numerically higher for SCZ, variance effects within the visual network and (**J-L**) the sensory-motor network were not significant for any *Group* main effects or interactions. Also between-group t-tests prior to GSR were not significant (both p-values >0.23). Note: Across the bar plot panels we highlight the pairwise effects (only significant for SCZ relative to HCS in panel B & E) as well as all the *Group x Processing* (only significant for SCZ vs. HCS in the higher-order associative networks, panels B & E). These preferential network effects further rule out movement concerns; it is highly unlikely that two separate patient samples collectivity move their fronto-parietal and default networks more than their visual-motor networks.

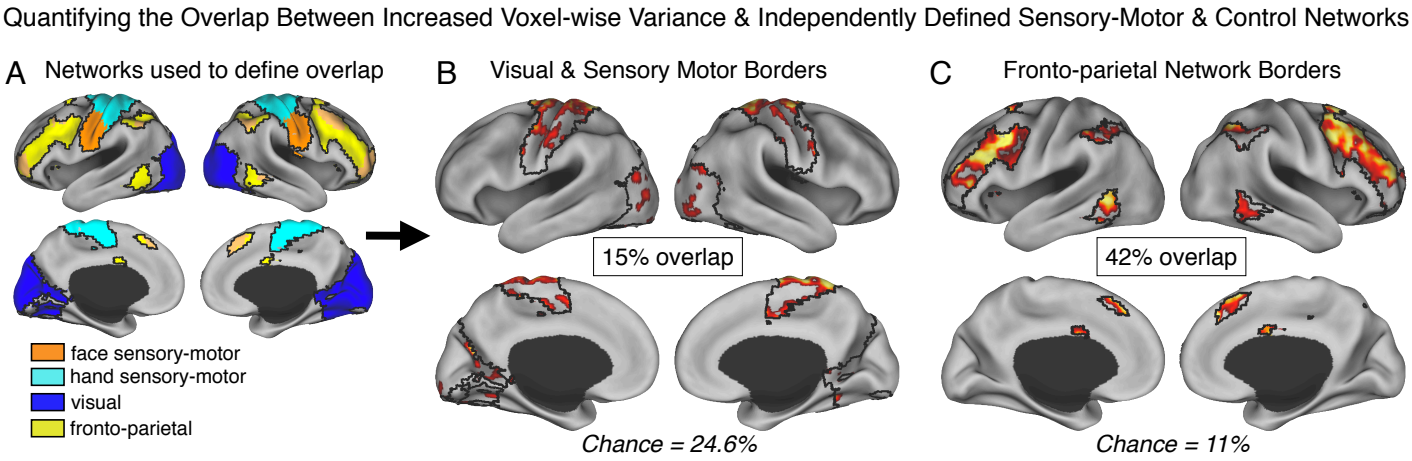

Figure S13. Characterizing Specificity of the Voxel-wise Effects in SCZ – Focus on Large Scale Networks. We discovered significantly elevated voxel-wise variability across distributed cortical and subcortical regions in SCZ (**Fig. 3**). The pattern appeared widespread and rather diffuse, but it raises the question of whether these effects co-localized preferentially within specific large-scale networks. To this end, we formally quantified the pattern of elevated voxel-wise variability with specific *a priori* defined networks, as done for the GS effects in **Fig. S12**. (**A**) To accomplish this, we used the sensory-motor network map defined independently via resting-state by Power and colleagues (2011) (obtained with permission from (45)). **(B)** 15% of all voxels belonging to sensory-motor networks defined *a priori* (black borders) overlapped with the elevated voxel-wise variance map (chance overlap = 24.6%). **(C)** In contrast, 42% of all voxels belonging to fronto-parietal networks defined *a priori* (black borders) overlapped with the elevated voxel-wise variance map (chance overlap = 11% because there are less total fronto-parietal voxels). To confirm this statistically we ran a binomial test for difference in proportions: 42% versus 15% of spatial overlap between the two sets of networks significantly exceeded the proportion difference expected by chance alone ($Z=35.51$, $p<0.000001$, binomial test for difference between proportions). Importantly, the fronto-parietal overlap (panel C) was significantly higher than expected by chance alone ($Z=65.94$, $p<0.000001$, binomial test for proportions), whereas the sensory-motor overlap was actually significantly lower than would have been expected by chance (as by chance alone we would expect approximately ~24.6% overlap (given the number of voxels in the sensory-motor networks defined here relative to all possible voxels belonging to the Power networks). Again, these preferential network effects further rule out movement concerns; it is highly unlikely that two separate patient samples collectively move their fronto-parietal and default networks more than their visual-motor networks.

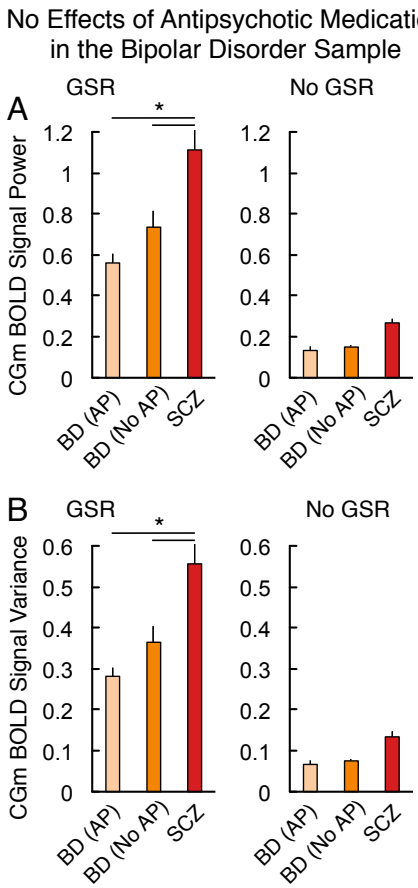

Figure S14. Examining Effects of Long-term Antipsychotic Treatment on Cortical Gray Matter (CGm) Variance Effects. One concern regarding present cross-diagnostic results pertains to cumulative effects of antipsychotic treatment over extended periods of time, which may differ across diagnostic groups. The cumulative medication impact is very difficult to quantify accurately, particularly the precise level and exact duration of compounding medication effects on brain function in SCZ. Here we attempted a proxy analysis that adds confidence that the core clinical SCZ effects are not confounded by medication. Specifically, we identified N=25 bipolar patients who were treated in conjunction with anti-psychotic medication during their illness course (BD AP). Next, we explicitly compared the CGm (**A**) power and (**B**) variance profiles between SCZ and bipolar patients that received long-term antipsychotic treatment (BD AP) as well as those bipolar patients that were not treated by antipsychotics (BD No AP). The core SCZ effects still remained evident across both comparisons for power (SCZ vs. BD-AP [$t(206)=2.16$, $p<.035$]; SCZ vs. BD-noAP [$t(183)=1.99$, $p<.05$]) as well as variance (SCZ vs. BD-AP [$t(206)=2.17$, $p<.032$]; SCZ vs. BD-noAP [$t(183)=2.05$, $p<.045$]) measures. Moreover, the effects indicated a marginally lower level of CGm power/variance for the BD AP patients. Put simply, those BD patients who received antipsychotic treatment exhibited marginally lower, not higher variance profiles – an effect that is inconsistent with the possibility that long-term anti-psychotic treatment alone is driving the core SCZ results. Nevertheless, longitudinal studies that carefully quantify the level of medication over time can further establish the impact of long-term medication dose/type on present effects.

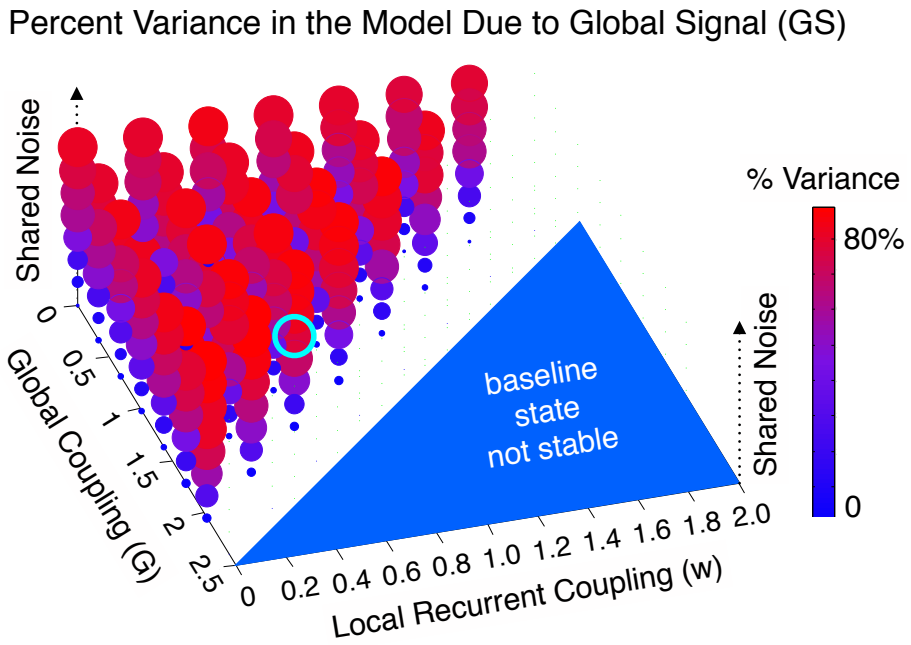

Figure S15. Examining Percentage of *in silico* Global Signal Variance as a Function of Model Parameters. As noted, Deco and colleagues (2013) utilized different default w/G/sigma values and did not incorporate a common shared signal component. Here we found that the original parameter choices, without a common shared signal, produced results that were somewhat unrealistic given the known large contribution of GS in empirical data (see **Fig. 1**). For instance, in our empirical data, we found that ~90% of the variance of the healthy control BOLD signal was accounted for by the GS component (although this likely collapses neuronal and physiological signal present *in vivo*). Therefore, we implemented a common noise input into our model architecture. Here we show a full parameter sweep of w (x-axis), G (y-axis), and magnitude of shared noise (z-axis) parameters and examine the percentage of signal variance that is attributed to the global signal (GS) component in the model. As the shared signal component increases (z-axis) we found that the percentage of signal variance attributable to the GS increased (color bar on the right; also marked as increasing circles). This was a generic property of the model irrespective of w/G parameter combinations; we selected a model regime where simulated neuronal GS accounted for 77% of the total simulated BOLD signal variance at default w/G values (cyan circle).

SI REFERENCES

1. Anticevic A*, et al.* (2013) Characterizing Thalamo-Cortical Disturbances in Schizophrenia and Bipolar Illness. *Cereb. Cortex* [Epub].

2. First MB, Spitzer RL, Miriam G, & Williams JBW (2002) *Structured clinical interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP)* (Biometrics Research, New York State Psychiatric Institute, New York).

3. Spreen O & Strauss E (1998) *A compendium of neuropsychological tests: Administration, norms, and commentary* (Oxford University Press, New York) 2nd Ed.

4. Lezak MD (1995) *Neuropsychological Assessment* (Oxford University Press, New York) 3rd Ed.

5. Barch DM & Ceaser A (2012) Cognition in schizophrenia: core psychological and neural mechanisms. *Trends In Cognitive Sciences* 16(1):27-34.

6. Krystal JH*, et al.* (2006) The vulnerability to alcohol and substance abuse in individuals diagnosed with schizophrenia. *Neurotox. Res.* 10:235-252.

7. Anticevic A*, et al.* (2012) Global Prefrontal and Fronto-amygdala Dysconnectivity in Bipolar I Disorder with Psychosis History. *Biol. Psychiatry* 73(6):565-573.

8. Anticevic A, Repovs G, Corlett PR, & Barch DM (2011) Negative and Non-emotional Interference with Visual Working Memory in Schizophrenia. *Biol. Psychiatry* 70(12):1159-1168.

9. Cole MW, Anticevic A, Repovs G, & Barch DM (2011) Variable global dysconnectivity and individual differences in schizophrenia. *Biol. Psychiatry* 70(1):43-50.

10. Glahn DC, Bearden CE, Bowden CL, & Soares JC (2006) Reduced educational attainment in bipolar disorder. *J. Affect. Disord.* 92:309–312.

11. Andreasen NC, Pressler M, Nopoulos P, Miller D, & Ho B-C (2010) Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biol. Psychiatry* 67(3):255-262.

12. Kay SR, Fiszbein A, & Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13:261-276.

13. Mayer AR*, et al.* (2012) Functional imaging of the hemodynamic sensory gating response in schizophrenia. *Hum. Brain Mapp.* 34(9):2302-2312.

14. Stephen JM*, et al.* (2013) Using joint ICA to link function and structure using MEG and DTI in schizophrenia. *Neuroimage* 83:418-430.

15. Hanlon FM*, et al.* (2011) Bilateral hippocampal dysfunction in schizophrenia. *Neuroimage* 58(1158-1168).

16. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, & Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59(3):2142-2154.

17. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, & Petersen SE (2012) Steps toward optimizing motion artifact removal in functional connectivity MRI; a reply to Carp. *Neuroimage* 76:439-441.

18. Anticevic A*, et al.* (2012) NMDA Receptor Function in Large-Scale Anti-Correlated Neural Systems with Implications for Cognition and Schizophrenia. *Proceedings of the National Academy of Science U.S.A.* 109(41):16720-16725.

19. Biswal BB*, et al.* (2010) Toward discovery science of human brain function. *Proc Natl Acad Sci USA* 107(10):4734-4739.

20. Fischl B, Salat DH, Busa E, Albert M, & Dieterich M (2002) Whole Brain Segmentation Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron* 33(3):341-355.

21. Repovs G, Csernansky JG, & Barch DM (2011) Brain network connectivity in individuals with schizophrenia and their siblings. *Biol. Psychiatry* 15(69):967-973.

22. Smith SM & Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1):83-98.

23. Fox MD*, et al.* (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102(27):9673-9678.

24. van Rossum G & Drake FL (2001) Python Reference Manual. ed Drake GvRaFL (PythonLabs, Virginia, USA).

25. Cole MW, Pathak S, & Schneider W (2010) Identifying the brain's most globally connected regions. *Neuroimage* 49(4):3132-3148.

26. Goldman-Rakic PS (1991) Prefrontal cortical dysfunction in schizophrenia: The relevance of working memory. *Psychopathology and the Brain*, eds Carroll BJ & Barrett JE (Raven Press, Ltd., New York), pp 1-23.

27. Saad ZS*, et al.* (2012) Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connectivity* 2(1):25-32.

28. Cohen AL*, et al.* (2008) Defining functional areas in individual human brains using resting functional connectivity MRI. *Neuroimage* 41(1):45-57.

29. Friston KJ (2011) Functional and effective connectivity: a review. *Brain Connectivity* 1(1):13-36.

30. Welsh RC, Chen AC, & Taylor SF (2010) Low-frequency BOLD fluctuations demonstrate altered thalamocortical connectivity in schizophrenia. *Schizophr. Bull.* 36(4):713-722.

31. Woodward ND, Karbasforoushan H, & Heckers S (2012) Thalamocortical dysconnectivity in schizophrenia. *Am. J. Psychiatry* 169(10):1092-1099.

32. Fischl B*, et al.* (2004) Sequence-independent segmentation of magnetic resonance images. *Neuroimage* 23 Suppl 1:S69-84.

33. Deco G*, et al.* (2013) Resting-State Functional Connectivity Emerges from Structurally and Dynamically Shaped Slow Linear Fluctuations. *J. Neurosci.* 33(27):11239-11252.

34. Wong KF & Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26(4):1314-1328.

35. Hagmann P*, et al.* (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol.* 6(7):e159.

36. Deco G & Jirsa VK (2012) Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* 32(10):3366-3375.

37. Liu YJ*, et al.* (2007) Are the local blood oxygen level-dependent (BOLD) signals caused by neural stimulation response dependent on global BOLD signals induced by hypercapnia in the functional MR imaging experiment? Experiments of long-duration hypercapnia and multilevel carbon dioxide concentration. *AJNR Am. J. Neuroradiol.* 28(6):1009-1014.

38. Phillips ML, Travis MJ, Fagiolini A, & Kupfer DJ (2008) Medication effects in neuroimaging studies of bipolar disorder. *Am. J. Psychiatry* 165(3):313-320.

39. Andreasen NC*, et al.* (1994) Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* 266:294-298.

40. Murphy K, Birn RM, Handwerker DA, Jones TB, & Bandettini PA (2009) The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44(3):893-905.

41. Power JD*, et al.* (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* 84(320-341).

42. Feinberg DA*, et al.* (2010) Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS ONE* 5(12):p. e15710.

43. Klingner CM*, et al.* (2013) Thalamocortical connectivity during resting state in schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* [Epub ahead of print].

44. Baker JT*, et al.* (2014) Disruption of cortical association networks in schizophrenia and psychotic bipolar disorder. *JAMA Psychiatry* 72(2):109-118.

45. Power JD*, et al.* (2011) Functional network organization of the human brain. *Neuron* 72(4):665-678.